

ASIMETRÍAS Y PERSISTENCIA

Tesis Doctoral

Autor: Oscar Martínez Ibáñez

Director: Jesús Gonzalo Muñoz



Departamento de Economía

UNIVERSIDAD CARLOS III DE MADRID

Getafe, Septiembre de 2004

A mis padres.

A Olalla

Índice general

1. Introducción	7
1.1. Estructura y principales aportaciones de la Tesis	12
1.2. Conceptos previos	15
1.2.1. Función respuesta a un impulso	17
1.2.2. Persistencia	19
1.3. Conclusiones	22
1.4. Bibliografía	24
2. Modelos ARTIMA. El Caso Observable.	29
2.1. Introducción	30
2.2. Modelos ARTIMA.	32
2.3. ARTIMA-Observable	34
2.3.1. Función Respuesta a un Impulso	37
2.3.2. Invertibilidad	47
2.4. Estimación	50
2.5. Inferencia	52
2.5.1. Inferencia cuando r^0 es conocido	53
2.5.2. Inferencia cuando r^0 es desconocido	56
2.5.3. Comportamiento de los contrastes en muestras finitas	60

2.6.	Aplicación a la tasa de desempleo.	62
2.7.	Conclusiones	70
2.8.	Bibliografía	72
2.9.	Anexo 1	76
2.10.	Anexo 2	81
2.10.1.	Consistencia y tasa de convergencia.	82
2.10.2.	Inferencia.	103
3.	Modelos ARTIMA-Shock. El Caso No Observable.	125
3.1.	Introducción	126
3.2.	Modelo ARTIMA-Shock	128
3.2.1.	Función Respuesta a un Impulso.	133
3.2.2.	Invertibilidad	136
3.3.	Estimación	141
3.4.	Inferencia	144
3.4.1.	Inferencia para r^0 conocido	144
3.4.2.	Inferencia cuando r^0 es desconocido	146
3.4.3.	Comportamiento de los contrastes en muestras finitas	150
3.5.	Aplicación al Índice de Precios Industriales	153
3.6.	Conclusión	159
3.7.	Bibliografía	162
3.8.	Anexo 1.	163
3.9.	Anexo 2.	166
3.9.1.	Consistencia y Tasa de Convergencia.	168
3.9.2.	Inferencia.	193

4. Método No Lineal de Descomposición en Transitorio y Permanente.	219
4.1. Introducción	220
4.2. Contraste de Existencia de Asimetrías	223
4.2.1. Comportamiento del contraste en muestras finitas	226
4.3. Descomposición Permanente-Transitorio	227
4.4. Aplicación a los Precios de Activos: Una medida de la Calidad del Mercado de Activos	233
4.5. Aplicación al PNB: Persistencia Asimétrica de las Perturbaciones a la Pro- ducción	238
4.6. Conclusión	245
4.7. Bibliografía	246
4.8. Anexo	249
5. Modelos Threshold y Larga Memoria	261
5.1. Introducción	262
5.2. Definiciones de Larga Memoria	264
5.3. Modelos Threshold y Larga Memoria	265
5.4. Estudio de Monte Carlo	271
5.5. Conclusiones	288
5.6. Bibliografía	289
6. Consideraciones finales y líneas de investigación futura.	291
6.1. Bibliografía	301

Agradecimientos

Muchas han sido las personas que con su apoyo y dedicación han hecho posible que esta Tesis haya llegado, por fin, a puerto.

En primer lugar me gustaría agradecer a mi director de tesis, Jesús Gonzalo, su confianza, apoyo, paciencia y tiempo dedicado. Con él he aprendido muchas cosas, no todas recogidas en esta Tesis, y por ello, le estaré siempre agradecido.

En segundo lugar me gustaría expresar mi gratitud a los miembros de los departamentos de Estadística y Economía de la Universidad Carlos III, por su confianza y por haber puesto a mi disposición todos los medios necesarios para la elaboración de esta Tesis. En especial, ha sido una suerte poder contar con los conocimientos y ayuda de Esther Ruiz, Olivier Nuñez, Paco Marmol, Miguel Delgado, Carlos Velasco y Juan José Dolado. Al profesor David Hendry tengo que agradecerle su disponibilidad y cordial acogida durante mi estancia en la Universidad de Oxford.

Además he tenido la suerte de compartir estos años de doctorado con estupendos compañeros. Como Omar, Stefano y Amparo, cuyo apoyo informático, tan útil fue en esta fase final. De forma especial, quiero recordar, el tiempo pasado con Jose Olmo, Sara, Alfredo Chourraut, Neida y Lulo, los menús con Ramirin, la ayuda tantas veces brindada por Juan Txarles, los cursos de guitarra y el cariño de Nieves, la dulce amistad de Natalia, Rusiña, Luana o Virginia, las charlas y lo que me han hecho reír Pablo Goñi, Rocio, Angelines, Txarly, Paquito o Pepe, y por supuesto, la paciencia y calor de Andrés, compañero de viaje, de piso y de despacho. A todos ellos les debo su apoyo constante. Su amistad ha hecho que esta empresa haya merecido la pena.

Con Javi, Fernando y Ana compartí la infancia, la juventud y espero que la vejez. Gracias por vuestra sincera amistad.

También me gustaría recordar a Itziar, Jose Mari, Nagore y Haritz, por tratarme siempre también y hacerme sentir como de la familia.

Gracias a prima Olvido y a Josema, por su amistad y apoyo en los peores momentos. Nunca os podré devolver todo lo que me habéis dado. Gracias Olalla, por tu comprensión, por tu presencia, por tus abrazos, por tu cariño. Sin ti, esta tesis no hubiera sido posible. Tampoco hubiera sido posible sin el apoyo de mis hermanos, Mario y Miriam, y de mis padres, a ellos les debo todo. Gracias por vuestro sacrificio y por vuestro cariño.

Capítulo 1

Introducción

La sucesión de los valores observados de una determinada variable, recogidos a lo largo del tiempo, es lo que se conoce como serie temporal. En economía, las series temporales suelen verse como realizaciones de procesos estocásticos, dada la existencia de perturbaciones o innovaciones inesperadas. La identificación de estas perturbaciones y su dinámica o ley de probabilidad se revela como un aspecto fundamental en la teoría económica, ya que serán la base para la validación de teorías y la compresión y predicción de la variable que conforma la serie. Una parte importante de este análisis consiste en la construcción y estimación de modelos de series temporales empíricos que capturen las posibles dinámicas de las series. En la práctica, las dinámicas suelen ser excesivamente complejas, con lo que su estudio se debe centrar en aquellas características que el investigador considera fundamentales, como la persistencia, volatilidad, asimetrías etc.

En este contexto, los modelos lineales han dominado el desarrollo de las series temporales durante mucho tiempo. Este dominio se ha basado en la metodología propuesta por Box y Jenkins (1976) sobre los modelos autoregresivos integrados de medias móviles, ARIMA. Las principales ventajas de estos modelos son su completo desarrollo teórico y su relativo éxito como instrumento para el análisis, predicción y control en las series temporales. La

fuerza de los modelos ARIMA reside en el Teorema de representación de Wold, (pág. 187 de Brockwell y Davis, 1993) que garantiza que el componente estocástico de cualquier proceso estacionario puede ser aproximado por un modelo ARMA. Sin embargo, la existencia de una determinada modelización para una serie no garantiza que ésta sea la mas adecuada para estudiar todas las características de la misma.

Este es el caso de muchas teorías económicas que suelen implicar características no estándar o no lineales en las variables, como son los casos de Business Cycle, costes de transacción, puntos de saturación, diferentes políticas de almacenamiento o “Buffer stock”, funciones de inversión, producción o curvas de Phillips no lineales, etc. Estas características no lineales suelen incluir, por ejemplo, no normalidad, asimetrías, bimodalidad, relaciones no lineales entre las variables, irreversibilidad, sensibilidad a las condiciones iniciales, etc. Dichas características están fuera del alcance de los modelos gaussianos. Por lo tanto, en un marco estacionario podemos definir los modelos no lineales como aquellos para los que la representación de Wold no genera innovaciones independientes. Entre los modelos no lineales podemos distinguir dos tipos. Unos, los que se denominan “implícitos”, mantienen la estructura lineal de los modelos ARMA, pero las innovaciones poseen una distribución no gaussiana apropiada, ver Tong (1990). Sin embargo, como recoge Potter (1994), los modelos no lineales “implícitos” no podrán generar dinámicas asimétricas, que sólo podrán ser recogidas a través de los modelos no lineales “explícitos”, en los que la variable de interés se expresa como una función no lineal de su pasado. Son este tipo de modelos en los que nos centraremos a lo largo de la Tesis. El interés en estudiar las características no lineales de las series temporales, ha hecho que en los últimos años haya habido un desarrollo muy importante en la modelización no lineal de series temporales, véase, entre otros, Barnett et al. (2000), Day and Chen (1993), Fan y Yao (2003), Pesaran y Potter (1993), Granger and Teräsvirta (1993), Tong (1990). Lamentablemente, los modelos no lineales propuestos por la teoría económica distan de ser exactos. Esto, unido a la infinidad de posibles modelos

no lineales y la complejidad de un desarrollo teórico completo, ha impedido que su avance haya sido todo lo amplio que cabría desear. Por ello, la modelización econométrica debe enfatizar modelos relativamente simples y aplicables, que por supuesto, sean capaces de recoger las características fundamentales de la serie que se quiere estudiar.

Entre los modelos no lineales más extendidos, cabe destacar los modelos threshold. La clave fundamental de los modelos threshold es que se trata de aproximaciones lineales por regímenes. Además de poseer una modelización e interpretación simple e intuitiva, cabe destacar su versatilidad. Prácticamente, cualquier función suave no lineal puede ser aproximada, tanto como deseemos, por un modelo threshold. Es lo que se denomina “principio threshold”, propuesto por Tong (1987, 1990). Estos modelos se han usado con relativo éxito en el estudio del “Business Cycle”, la ley de precio único (PPP), el tipo de cambio, en el estudio de volatilidades, etc.

El estudio de las asimetrías y no linealidades en economía se inicia con las teorías del Business Cycle, descritas ya por Keynes (1936) o Burns y Mitchell (1946), según las cuales, las series económicas no se comportan igual en las expansiones que en las recesiones. Concretamente, se definen dos tipos de asimetrías: “Deepness”, que implica que la distancia a la tendencia es mayor en las recesiones que en las expansiones; y “steepness” que supone que las recesiones son mas abruptas que las expansiones. La literatura que estudia la modelización no lineal de estas asimetrías en el Business Cycle es muy extensa, véase por ejemplo, Acemoglu y Scott (1997), Geweke y Terui (1991), Hamilton (1989), Neftçi (1984), Potter (1991), Semmler (1994), Thygesen et al. (1991). Igualmente los modelos no lineales se han utilizado con éxito en el estudio de la volatilidad, como es el caso de los modelos autoregresivos de heterocedasticidad condicional (ARCH), introducidos por Engle (1982). Estos modelos aunque lineales en la media, no son completamente lineales por su modelización de la varianza. Posteriormente se han introducido estructuras no lineales más complejas, que permiten comportamientos asimétricos más completos de la volatilidad.

Otra característica relevante y común en las series económicas es la persistencia cuyo estudio a través de modelos no lineales ha sido menos desarrollado. En el caso de la modelización lineal ARIMA, la persistencia se recoge a través de las raíces unitarias, que implican que todas las perturbaciones provocan un efecto permanente y del mismo grado en la serie. La ausencia de raíces unitarias supone que todas las perturbaciones son transitorias. Posteriormente, estos modelos se extendieron a través de los modelos fraccionalmente integrados, ARFIMA, que permiten que las perturbaciones tengan un mayor grado de persistencia, sin llegar a ser permanentes. En todos los casos, todas las perturbaciones de la serie tendrán el mismo grado de persistencia, lo que en un contexto económico supone una restricción muy fuerte. Las asimetrías en la persistencia de las perturbaciones fueron inicialmente estudiadas por Wecker (1981), que introduce un modelo de medias móviles con threshold, en el que la persistencia depende del signo de la perturbación. Posteriormente, Beaudry y Koop (1993) analizan el mismo tipo de asimetrías utilizando una variable proxy para el signo de la perturbación. Este trabajo es revisado por Hess e Iwata (1997) y Elwood (1997). Este último, generaliza el modelo de Wecker al caso autoregresivo. Sin embargo, en ninguno de estos casos hay un estudio detallado de las propiedades asintóticas de los estimadores, y aunque existe asimetría en la persistencia, las perturbaciones son o todas transitorias o todas permanentes. Finalmente, dentro de los modelos no lineales suaves o “smooth transition model”, cabe destacar el trabajo de Engle y Smith (1999) en el que introducen los modelos que denominan “Stochastic Permanent Break”, (STOPBREAK), en los que la persistencia de la perturbación dependerá de su tamaño. De nuevo, a pesar de la diferente persistencia en las perturbaciones, todas ellas son permanentes.

El objetivo de esta Tesis es extender el estudio de las asimetrías en la persistencia a un contexto más general. Para ello, permitiremos que las perturbaciones que afectan a la serie puedan tener un distinto grado de persistencia dependiendo de determinadas propiedades de la perturbación o de alguna variable económica de interés. Un caso preferente y muy

relevante será la existencia de perturbaciones transitorias y permanentes, que hasta la fecha no ha sido tratado en este marco por la literatura. Como bien es sabido (ver Quah, 1992), la persistencia en el largo plazo de una perturbación no es suficiente para su identificación. Son necesarios supuestos adicionales que hasta la fecha no se han relacionado con las asimetrías (ver Harvey 1985, Clark 1987 y Pierce 1979, entre otros). En esta Tesis se muestra que las asimetrías en la persistencia de las perturbaciones, dependiendo de determinadas propiedades o características de una variable económica (que puede ser la propia perturbación), son suficientes para su identificación. Este análisis deriva en la propuesta de un nuevo método de identificación del componente permanente y transitorio de la serie más robusto y endógeno.

Para todo ello, se introducen los modelos que denominaremos autoregresivos integrados con threshold en la media móvil, ARTIMA. Estos modelos se caracterizan por poseer una raíz unitaria en el componente autorregresivo y diferentes regímenes en la media móvil, con la posibilidad de que alguno de ellos sea no invertible. En cada caso se estudiarán las condiciones suficientes para la invertibilidad del modelo, así como para la existencia de asimetrías en las perturbaciones en términos de transitorio y permanente. Igualmente, se estudiará la consistencia y distribución asintótica de los estimadores que permitan la inferencia estadística sobre los aspectos relevantes del modelo. Éstos, se traducen en dos hipótesis fundamentales para la teoría económica, que son, la existencia de asimetrías y la existencia de perturbaciones transitorias y permanentes. Finalmente, la tesis abordará un estudio detallado de la relación entre modelos con este tipo de asimetrías en la persistencia y los modelos de larga memoria.

1.1. Estructura y principales aportaciones de la Tesis

La presente Tesis Doctoral se estructura en seis capítulos: El primero es un capítulo introductorio, los Capítulos 2, 3, 4 y 5 constituyen el núcleo central de la Tesis y recogen las principales aportaciones novedosas de la misma. En el sexto se muestran las conclusiones generales del presente trabajo.

El Capítulo 1 introduce y expone la motivación general de la Tesis. La Sección 1.2 introduce la función respuesta a un impulso, además discute algunos conceptos relevantes que se basan en dicha función, los cuales se usarán de forma continua en el resto de los capítulos, como son los conceptos de persistencia, perturbación permanente y perturbación transitoria.

El Capítulo 2 introduce un nuevo tipo de modelos no lineales que denominaremos modelos ARTIMA. Estos modelos facilitarán el estudio de un nuevo tipo de asimetría en la persistencia en términos de transitorio y permanente. Los modelos ARTIMA son procesos integrados con una modelización threshold o por umbrales de su componente de medias móviles, lo que implica diferentes regímenes lineales en dicho componente. La aportación fundamental de los modelos ARTIMA es la posibilidad de que las perturbaciones pertenecientes a determinados regímenes sean netamente transitorias, mientras que las del resto, son permanentes. Dependiendo de las propiedades de la variable que determine el régimen al que pertenece la perturbación, se distinguirán dos clases fundamentales de modelos ARTIMA, los modelos ARTIMA-Observables, cuando esta variable es observable, y los modelos ARTIMA-Shock, cuando esta variable es la propia perturbación, y por tanto, no observable. En este capítulo también se estudiarán en detalle las propiedades de los modelos ARTIMA-Observables, fundamentalmente las relacionadas con su invertibilidad y persistencia. Se propone el método de mínimos cuadrados condicionales para la estimación, obteniéndose las condiciones suficientes para la identificabilidad de los parámetros y la consistencia de los estimadores. Igualmente, se propone una estrategia para el contraste

de las hipótesis básicas del modelo, que serán dos: la primera, la existencia de asimetrías en la persistencia; la segunda, la existencia de perturbaciones transitorias y permanentes. Finalmente se obtendrá la distribución asintótica de los estadísticos propuestos para el contraste de estas hipótesis básicas, realizándose un pequeño experimento de Monte Carlo sobre su comportamiento en muestras finitas. El capítulo se cierra con una aplicación del modelo ARTIMA-Observable a la tasa de desempleo en Estados Unidos.

El Capítulo 3 estudiará en detalle los modelos ARTIMA-Shock introducidos en el Capítulo 2, y en los que la asimetría en la persistencia de las perturbaciones viene provocada por alguna propiedad de la propia perturbación. En este capítulo nos centraremos en las propiedades del tamaño y del signo. El caso del signo de la perturbación como causante de la asimetría en persistencia fue modelizado por Wecker (1981), aunque hasta la fecha no se han obtenido resultados de teoría asintótica para los métodos de estimación tradicionales. El caso del tamaño, fue introducido por Engle y Smith (1999), aunque en el modelo que proponen no se admite la posibilidad de asimetrías en términos de perturbaciones transitorias y permanentes. Así, los modelos ARTIMA-Shock desarrollados en este capítulo suponen una generalización de las características de la asimetría, permitiendo la existencia de perturbaciones transitorias y permanentes. Otra aportación fundamental de este capítulo es la obtención de resultados asintóticos para los métodos de estimación y contraste propuestos. Como en el caso Observable, se propone la estimación por mínimos cuadrados condicionales para la que se obtiene su comportamiento asintótico. Éste supondrá una novedad respecto a los resultados tradicionales en la estimación threshold. Cuando la variable threshold es la propia perturbación y el modelo no es continuo, todos los estimadores, y no solo el del parámetro threshold, serán superconsistentes. Igualmente, se muestra la distribución asintótica de los estadísticos propuestos para el contraste de las hipótesis básicas del modelo, existencia de asimetrías y existencia de perturbaciones transitorias. Como en el capítulo anterior, también se estudiarán las propiedades básicas

del modelo, invertibilidad y persistencia, y el comportamiento en muestras finitas de los contrastes propuestos. Finalmente, el capítulo se cierra con la aplicación de los resultados obtenidos para el caso del signo a diferentes series de índices de precios industriales utilizadas por Wecker.

El Capítulo 4 es una aplicación teórica del nuevo método de identificación de las perturbaciones introducido con los modelos ARTIMA. Este nuevo método de identificación presenta dos características fundamentales. La primera, hace referencia a la posibilidad de identificar las perturbaciones en base a criterios económicos y lógicos. La segunda, a la posibilidad de contrastar la validez de dichos criterios. Basado en este nuevo método de identificación, este capítulo presenta una nueva definición del componente transitorio y permanente de una serie. Se analizarán las ventajas que presenta la nueva definición frente a las ya existentes. Igualmente, se estudiará el efecto que la modelización ARTIMA tiene en los métodos de descomposición tradicionales, fundamentalmente los propuestos por Beveridge y Nelson (1981) y Componentes Inobservados (UC). El Capítulo finaliza con dos aplicaciones de los métodos propuestos a los precios de los activos y al Producto Nacional Bruto Real de Estados Unidos. En el primer caso, la descomposición permanente-transitorio sirve para medir la eficiencia del mercado. En el segundo caso, la descomposición permitirá estudiar en detalle los ciclos económicos.

El Capítulo 5 analiza la relación existente entre los modelos no lineales threshold y la propiedad de larga memoria, que tradicionalmente se ha modelizado a través de los procesos fraccionalmente integrados o ARFIMA. En concreto se analizarán los modelos ARTIMA, introducidos a lo largo de la Tesis, y los modelos Threshold Unit Root, TUR, introducidos por González y Gonzalo (1998). Este tipo de relación ha sido estudiada para otro tipo de modelos no lineales por Diebold e Innoue (2001), Granger y Hyung (1999) y Granger y Terasvirta (1999) entre otros. En este capítulo se prueba que esta relación es similar a la existente entre la propiedad de larga memoria y los modelos con break aleatorios en la

media, utilizados por la teoría para justificar la existencia de larga memoria en las series económicas. Esta relación depende del tamaño muestral, por ello, a través de un estudio de Monte Carlo se muestra el comportamiento del estimador del parámetro de larga memoria en los modelos ARTIMA y TUR para distintos tamaños y bajo diferentes condiciones. También se estudia el comportamiento de los contrastes de larga memoria propuestos por Geweke y Porter-Hudak (1983) y Tanaka (1999), y el contraste de raíz unitaria propuesto por Dickey y Fuller (1979, 1981). Finalmente, este comportamiento se compara con el que presentan dichos métodos en los modelos lineales ARIMA y ARFIMA derivados de los modelos no lineales ARTIMA y TUR. El objetivo de esta comparación es comprobar que porcentaje de esta larga memoria se debe netamente a las asimetrías.

Finalmente, el Capítulo 6 expone las líneas de investigación futuras y presenta las principales conclusiones de la Tesis.

1.2. Conceptos previos

El objeto de la Tesis es el estudio detallado de las asimetrías en persistencia, así como de los modelos propuestos para su estudio. Por lo tanto, a lo largo de la misma se hará un uso continuado del concepto de persistencia. Este concepto ha sido ampliamente tratado en la literatura, aunque las definiciones que se han proporcionado de él no siempre han sido coincidentes, sobre todo en el contexto de modelos no lineales. Por ello es conveniente establecer desde el principio los matices que se consideran en esta Tesis.

Lo primero que conviene hacer notar es que, en un marco de asimetrías en persistencia donde las perturbaciones pueden ser permanentes o transitorias en función de las propiedades de alguna variable relevante, la persistencia de la serie no va a coincidir con la persistencia de todas las perturbaciones que la afectan. En este sentido, a lo largo de la Tesis nos centraremos en la persistencia de las perturbaciones y no de la serie.

Empezamos definiendo el tipo de perturbaciones que vamos a estudiar. Éstas serán las que se producen a través de las innovaciones ξ_t que definimos como la siguiente diferencia de martingalas (mds) :

$$\xi_t = y_t - E[y_t | \mathfrak{F}_{t-1}],$$

donde y_t es la serie de interés, $\mathfrak{F}_{t-1} = \sigma(y_{t-1}, y_{t-2}, \dots)$ el sigma álgebra de Borel generado por el pasado de la serie y_t hasta $t - 1$ y $E[y_t | \mathfrak{F}_{t-1}]$ la esperanza condicional. La selección de esta definición se basa en su fácil interpretación y su relación con las expectativas racionales. Una perturbación es aquello que se produce en t y que no era esperado en $t - 1$. Otra posibilidad, como proponen Gallant et al. (1993) en un contexto no paramétrico, es definir las perturbaciones en términos de variaciones de la propia serie y_t , expresadas como δy_t . Esta definición presenta la dificultad de identificación e interpretación de δ .

El concepto de persistencia de una perturbación en una serie temporal se refiere a la existencia de una dependencia entre las observaciones futuras y la propia perturbación. Es importante destacar que puede hacer referencia tanto a la duración de la relación como a la magnitud de la misma. En concreto, a lo largo de esta Tesis, las asimetrías harán referencia a las diferentes magnitudes en la relación entre las observaciones futuras y las perturbaciones cuando la distancia entre ambas tiende a infinito. Así, definiremos como perturbaciones transitorias aquellas que no tengan relación, magnitud cero, con las observaciones futuras cuando la distancia tienda a infinito. En cualquier otro caso, la perturbación será permanente. Por lo tanto, para definir la persistencia será necesaria una herramienta capaz de medir dicha relación. Esta herramienta es la función respuesta a un impulso. En la siguiente sección se hace un estudio detallado de las distintas definiciones de función respuesta a un impulso, así como de sus principales propiedades, centrándonos en el caso no lineal.

1.2.1. Función respuesta a un impulso

Las definiciones que se muestran en esta sección han sido adaptadas de los artículos de Koop et al. (1996) y Potter (2000). La función respuesta a un impulso mide el efecto de una perturbación en el comportamiento futuro de la serie. Por supuesto, hay muchas maneras de medir este efecto. Para el caso lineal distinguimos las siguientes cuatro funciones. Sea y_t la variable de interés y $L_k(y_t, y_{t-1}, \dots)$ el predictor lineal de y_{t+k} dada la realización (y_t, y_{t-1}, \dots) . Por simplicidad, asumiremos que tomadas d diferencias de y_t el proceso es estacionario, de forma que usando la representación de Wold tenemos que

$$L_k(y_t, y_{t-1}, \dots) = (t+k)\mu + \sum_{i=0}^{\infty} \psi_{k+i} u_{t-i}$$

donde u_t es un ruido blanco. Las diferentes definiciones para la función respuesta a un impulso son,

- La función respuesta a un impulso lineal, $(lir f_k(\delta))$ que mide el efecto sobre la predicción lineal de y_{t+k} de un cambio en y_t de tamaño δ :

$$lir f_k(\delta) = L_k(y_t + \delta, y_{t-1}, \dots) - L_k(y_t, y_{t-1}, \dots) = \psi_k \delta$$

- La derivada del predictor lineal con respecto a y_t :

$$dlir f_k = \frac{\partial L_k(y_t, y_{t-1}, \dots)}{\partial y_t} = \psi_k$$

- La función adaptativa lineal, que mide el cambio en la predicción lineal ante la nueva llegada de información en t :

$$lud_k(u_t) = L_k(y_t, y_{t-1}, \dots) - L_k(y_{t-1}, y_{t-2}, \dots) = \psi_k u_t$$

Estas tres definiciones comparten las siguientes propiedades; no dependen ni del futuro ni del pasado de la serie y son simétricas respecto a cualquier característica de la perturbación, esto es, el efecto de δ es igual al de $-\delta$ con signo opuesto y la mitad que 2δ . Además,

como es fácil comprobar, convenientemente estandarizadas las tres definiciones coinciden. Estas propiedades son fácilmente derivables si tenemos en cuenta las propiedades de la última definición de función respuesta a un impulso propuesta por Potter (2000) y que denomina “sample path operator”.

- Consideremos dos realizaciones de un ruido blanco para las innovaciones $\{u_t\}$, $\{u_j^*\}_{j=-\infty}^{\infty}$ y $\{u_j\}_{j=-\infty}^{\infty}$, iguales excepto en t . Podemos obtener, por lo tanto, $\{y_j^*\}_{j=-\infty}^{\infty}$ y $\{y_j\}_{j=-\infty}^{\infty}$. El “sample path operator” vendrá dado por

$$y_{t+k}^* - y_{t+k} = \psi_k [u_t^* - u_t]$$

Por lo tanto, en el caso de procesos lineales, el efecto en la serie de u_t^* respecto a u_t , solo depende de la diferencia entre ambas perturbaciones, no depende ni de u_t o u_t^* , ni del pasado o futuro de la serie.

Sin embargo, cuando el proceso es no lineal, el “sample path operator” dependerá de las perturbaciones comparadas, u_t^* y u_t , así como del pasado y futuro de la serie, con lo que las diferentes definiciones no coincidirán. Además, esta dependencia del “sample path operator” en las perturbaciones, el futuro o el pasado de la serie, invalida la representación de Wold de un modelo no lineal para estudiar las asimetrías de la serie (véase Potter, 2000).

Para extender las definiciones anteriores al caso no lineal Potter propone utilizar la esperanza condicional como predictor, por lo tanto, las extensiones se definirían como:

- $nlirf_k(\delta, \mathfrak{S}_t) = E[y_{t+k} | y_t + \delta, y_{t-1}, \dots] - E[y_{t+k} | y_t, y_{t-1}, \dots]$
- $dnlirf_k(\mathfrak{S}_t) = \lim_{\delta \rightarrow 0} \frac{\partial nlirf_k(\delta, \mathfrak{S}_t)}{\partial \delta}$
- $nlud_k(\mathfrak{S}_t) = E[y_{t+k} | \mathfrak{S}_t] - E[y_{t+k} | \mathfrak{S}_{t-1}]$

De esta forma, el efecto de la perturbación en el futuro de la serie será función de la propia perturbación (excepto en el caso de $dnlirf$) y del pasado de la serie. La dependencia

del futuro es resuelta considerando su valor esperado. Dado que las definiciones no han de coincidir, debemos decidir cual de ellas usaremos en la definición de persistencia. En el caso de la función $nlirf$, existe el problema de la selección e interpretación de δ . El problema fundamental de la definición basada en la derivada, $dnlirf$, es que no siempre está definida e impone el análisis de perturbaciones infinitesimales. Finalmente, teniendo en cuenta el tipo de perturbaciones objeto de estudio a lo largo de la Tesis, $y_t - E[y_t | \mathfrak{S}_{t-1}]$, la función $nlud$ se revela como la elección natural.

Por lo tanto, como hacen Potter y Koop et al., la definición de persistencia que usaremos a lo largo de la Tesis se basará en la función $nlud$, y que estos autores denominan, función respuesta a un impulso generalizada (GI), que tendrá la siguiente expresión,

$$GI(k, \xi_t, \mathfrak{S}_{t-1}) = E[y_{t+k} | \xi_t, \mathfrak{S}_{t-1}] - E[y_{t+k} | \mathfrak{S}_{t-1}],$$

cuando las innovaciones ξ_t sean la única información adicional de $t - 1$ a t . La función respuesta a un impulso dependerá, en general, del horizonte k , la perturbación, ξ_t , y el pasado \mathfrak{S}_{t-1} .

1.2.2. Persistencia

En el caso de la función respuesta a un impulso lineal, una perturbación será transitoria si los valores futuros de la serie no dependen de dicha perturbación, lo cual se cumplirá si y sólo si

$$\lim_{k \rightarrow \infty} dlirf_k = 0,$$

donde $dlirf$ puede ser sustituida por el resto de definiciones, dado que todas son equivalentes. Cuando el valor de la función en el límite sea distinto de 0, diremos que la perturbación es permanente. Respecto a la comparación de dos perturbaciones permanentes, diremos que es más persistente aquella para la que $\lim_{k \rightarrow \infty} dlirf_k$ es mayor.

En el caso no lineal, como ya hemos comentado, usar la esperanza incondicional de la GI puede eliminar efectos asimétricos de las perturbaciones, por lo tanto, lo conveniente será considerar la GI como una variable aleatoria en ξ_t o \mathfrak{S}_{t-1} (con ε_t y w_t realizaciones de dichas variables aleatorias (v.a)). En general, podremos condicionar en una determinada perturbación, ε_t , y considerar la GI como una v.a en la historia de la serie,

$$GI(k, \varepsilon_t, \mathfrak{S}_{t-1}) = E[y_{t+k} | \varepsilon_t, \mathfrak{S}_{t-1}] - E[y_{t+k} | \mathfrak{S}_{t-1}].$$

Alternativamente, podemos condicionar en una particular realización del pasado, $w_{t-1} \in \mathfrak{S}_{t-1}$, y considerar la GI como una v.a en la perturbación,

$$GI(k, \xi_t, w_{t-1}) = E[y_{t+k} | \xi_t, w_{t-1}] - E[y_{t+k} | w_{t-1}].$$

En la práctica, dependiendo de la información a priori y del análisis a realizar, lo habitual será condicionar en subconjuntos particulares de la historia o la perturbación,

$$GI(k, A, B) = E[y_{t+k} | \xi_t \in A, \mathfrak{S}_{t-1} \in B] - E[y_{t+k} | \mathfrak{S}_{t-1} \in B].$$

Así, si pensamos en y_t como el Producto Nacional Bruto (PNB), A podría ser el conjunto de las perturbaciones negativas y B el conjunto de todas las historias en las que la economía se encontraba en recesión en el periodo más reciente.

Al considerar la GI como una posible v.a, las medidas de persistencia de la respuesta se harán a través de la función de densidad de la GI. En concreto, y como propone Potter (2000), usaremos la dispersión de la GI como medida de persistencia. Bajo estacionariedad, ambas esperanzas condicionadas en la GI convergen al mismo valor, y por lo tanto, la GI y su dispersión convergerán a 0. Por el contrario, en el caso de un paseo aleatorio, la dispersión de la GI se mantendrá constante a medida que el horizonte de “predicción” se incrementa. En general, una mayor dispersión de la GI supone una mayor probabilidad de valores extremos, y por lo tanto, una mayor magnitud en la relación entre la observación futura y la perturbación actual.

Concretamente, diremos que una perturbación perteneciente al subconjunto A y condicionada a que $\mathfrak{S}_{t-1} \in B$ es transitoria, si la $GI(k, A, B)$ converge a cero cuando k tiende a ∞ con probabilidad uno. En cualquier otro caso, la perturbación perteneciente a A y condicionada al suceso $\mathfrak{S}_{t-1} \in B$ será permanente. Cuando no exista una fórmula cerrada para la GI, será difícil comprobar la propiedad de transitoriedad, dado que las esperanzas condicionadas se tendrán que estimar, y por lo tanto, presentarán siempre una varianza residual. Para comparar la persistencia de las perturbaciones es necesario, por tanto, definir una medida de dispersión. En el caso de que las densidades de la GI tengan la misma esperanza, Koop et al. (1996) y Potter (2000) proponen como medida de dispersión la dominancia estocástica de segundo orden. En concreto, la v.a X domina a Y en dominancia estocástica de segundo orden si,

$$\int_{\lambda < t} F_X(\lambda) \partial \lambda - \int_{\lambda < t} F_Y(\lambda) \partial \lambda \leq 0, \quad \forall t \in \mathfrak{R}$$

donde F_X y F_Y son las funciones de distribución de X e Y respectivamente. En general, nos referiremos a dominancia “estricta”, esto es, para algún t la desigualdad es estricta y por lo tanto cuando digamos que X domina a Y , supondremos que Y no domina a X . Así, si la persistencia disminuye a medida que el horizonte temporal se incrementa, la GI para el largo plazo debe dominar en dominancia estocástica de segundo orden a la de corto plazo. Si existe un efecto multiplicativo, esto es, la persistencia es cada vez mayor, la relación de dominancia será la opuesta. Igualmente, una perturbación perteneciente a A y condicionada a la historia B será más persistente que una perteneciente a A' condicionada a la historia B' en el horizonte k si $GI(k, A', B')$ domina a $GI(k, A, B)$ en dominancia estocástica de segundo orden.

El problema fundamental de la dominancia estocástica de segundo orden es que no genera un orden completo. Otra posibilidad sería usar la desviación típica como medida de dispersión, ya que ésta si generaría un orden completo, aunque lo haría a costa de una definición de persistencia más débil.

Finalmente, conviene comentar el caso particular en el que la $GI(k, A, B)$ converge a una constante $\mu_{A,B}$ distinta de 0 cuando $k \rightarrow \infty$. En este caso, la dispersión de $GI(k, A, B)$ se anula cuando k tiende a infinito, aunque hay un efecto permanente en la esperanza condicional. En este sentido, las perturbaciones pertenecientes a A condicionadas a la historia B , pueden considerarse permanentes. Este efecto permanente se produce en el componente determinista de la serie, por eso la $GI(k, A, B)$ converge a una constante. Esto es, la perturbación desaparecerá de la serie en el largo plazo, pero no su efecto en el componente determinista. Por otro lado, si consideramos la descomposición de Beveridge y Nelson (1981), (B-N en adelante), para el componente permanente y transitorio de la serie, las perturbaciones pertenecientes a A no formarán parte de este componente permanente. Esto se debe a que en la descomposición de B-N se eliminan del componente permanente los efectos deterministas de largo plazo. En este sentido, podríamos considerar que las perturbaciones pertenecientes a A son transitorias si la dispersión de $GI(k, A, B)$ tiende a cero, aunque su media sea distinta de cero. Sin embargo, en este caso la perturbación no podrá ser considerada puramente transitoria, ya que presenta un efecto permanente, aunque sea sólo en el componente determinista.

1.3. Conclusiones

Este primer capítulo introductorio ha pretendido, por una parte, plantear la motivación para escoger el tema de investigación que se desarrolla en la Tesis. Ésta fue doble, en primer lugar, la necesidad de identificar las perturbaciones en base a métodos endógenos y más robustos, para lo cual, se introduce una nueva clase de modelos que desarrollan un concepto de asimetría más general y en la que se basará la identificación de dichas perturbaciones. En segundo lugar, el desarrollo de la teoría asintótica necesaria para la estimación y el contraste de las hipótesis básicas de este modelo. Por otra parte, este capítulo presenta la

estructura y principales aportaciones, tanto en el campo teórico, como en su aplicación a series económicas reales. Por último, se ha introducido también los conceptos básicos en los que se basa el desarrollo teórico de la presente Tesis Doctoral.

1.4. Bibliografía

Acemoglu, D. y Scott, A. (1997): “Asymmetric business cycles: Theory and time-series evidence”. *Journal of Monetary Economics*, Vol. **40**, N3, 501-533.

Barnet, W.A., Hendry, D.F., Hylleberg, S., Teräsvirta, T., *Tøstheim*, D. y Würtz, A. (2000): *Nonlinear Econometric Modeling in Time Series Analysis*, Cambridge University Press.

Beaudry, P. y Koop, G. (1993): “Do recessions permanently change output?”. *Journal of Monetary Economics*, Vol. **31**, 149-163.

Beveridge, S. y Nelson, C.R. (1981): “A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘Business Cycle’ ”. *Journal of Monetary Economics*, Vol. **7**, 151-174.

Box, G.E.P. y Jenkins, G.M. (1976): *Time Series Analysis: Forecasting and Control*, Holden-Day.

Brockwell, P.J., Davies, R.A. (1993): *Time Series: Theory and Methods*, 2nd ed. Springer-Verlag, New York, NY.

Burns, A.F. y Mitchell, W.C. (1946): *Measuring Business Cycles*, Columbia University Press.

Clark, P.K., (1987): “The cyclical component of U.S. economic activity”. *Journal of Quarterly Journal of Economics*, Vol. **102**, 797-814.

Day, R.H. y Chen, P. (1993): *Nonlinear Dynamics and Evolutionary Economics*, Oxford University Press.

Dickey, D. y Fuller, W.A. (1979): “Distribution of the Estimators for Autoregressive Time Series With a Unit Root”. *Journal of the American Statistical Association*, Vol **74**, N 366, 427-431.

Dickey D.A. y Fuller W.A. (1981): "Likelihood ratio statistics for autoregressive time series with a unit root". *Econometrica*, Vol. **49**, 1057-72.

Diebold, F.X. y Inoue, A. (2001): "Long memory and regime switching". *Journal of Econometrics*, Vol. **105**, 131-159.

Elwood, S.K. (1998): "Is the persistence of shocks to output asymmetric?". *Journal of Monetary Economics*, Vol. **41**, 411-426.

Engle, R.F. (1982): "Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation". *Econometrica*, Vol. **50**, 987-1008.

Engle, R.F. y Smith, A.D. (1999): "Stochastic Permanent Breaks". *The Review of Economics and Statistics*, Vol. **81**, 553-574.

Fan, J. y Yao, Q. (2003): *Nonlinear Time Series. Nonparametric and Parametric Methods*, Springer-Verlag New York.

Gallant, A.R., Rossi, P.E. y Tauchen, G. (1993): "Nonlinear dynamics structures". *Econometrica*, Vol. **61**, 871-908.

Geweke J. y Porter-Hudak S. (1983): "The estimation and application of long memory time series models". *Journal of Time Series Analysis*, Vol **4**, N4, 221-238.

Geweke, J. y Terui, N. (1991): "Threshold Autoregressive Models for Macroeconomic Time Series: A Bayesian Approach". *Proceedings of the American Statistical Meeting, Business and Economics Section*, 42-50.

González, M. y Gonzalo, J. (1998): "Threshold Unit Root Processes". Working Paper.

Granger, C.W.J. y Hyung, N. (1999): "Occasional Structural Breaks and Long Memory". Working Paper, University of California - San Diego.

Granger, C.W.J. y Teräsvirta, T. (1993): *Modeling Nonlinear Economic Relationships*, Oxford University Press.

Granger, C.W.J., y T. Teräsvirta (1999): “A Simple Nonlinear Time Series Model with Misleading Linear Properties”. *Economic Letters*, Vol. **62**, 161-165.

Hamilton, J. (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle”. *Econometrica* Vol. **75**, N2, 357-384.

Harvey, A.C., (1985): “Trends and cycles in macroeconomics time series”. *Journal of Business and Economic Statistics*, Vol. **3**, 216-27.

Hess, G.D. e Iwata, S. (1997): “Asymmetric persistence in GDP? A deeper look at depth”. *Journal of Monetary Economics*, Vol. **40**, 535-554.

Keynes, J.M. (1936): *The General Theory of Employment, Interest and Money*, Macmillan.

Koop, G., Pesaran, M.H. y Potter, S.M. (1996): “Impulse response analysis in nonlinear multivariate models”. *Journal of Econometrics*, Vol. **74**, 119-147.

Neftci, S.N. (1984): “Are economic time series asymmetric over the business cycle?”. *Journal of Political Economy*, Vol. **92**, N 2, 307-328.

Pesaran, M.H. y Potter, S.M. (1993): *Nonlinear Dynamics Chaos and Econometrics*, John Wiley & Sons Ltd.

Pierce, D.A., (1979): “Signal extraction error in nonstationary time series”. *The Annals of Statistics*, Vol. **7**, No. 6, 1303-1320.

Potter, S. (1991): “A nonlinear approach to US GNP”,

Potter, S.M. (1994): “Asymmetric Economic Propagation Mechanisms”, *Business Cycles: Theory and Empirical Methods*. Kluwer Academic Publishers.

Potter, S.M. (2000): “Nonlinear impulse response functions”. *Journal of Economic Dynamics & Control*, Vol. **24**, 1425-1446.

Quah, D., (1992): “The relative importance of permanent and transitory components: identification and some theoretical bounds”. *Econometrica*, Vol. **60**, 107-118.

Semmler, W. (1994): *Business Cycles: Theory and Empirical Methods*, Kluwer Academic Publishers.

Tanaka, K. (1999): “The nonstationary fractional unit root”. *Econometric Theory*, Vol. **15**, N 4.

Thygesen, N., Kumaraswamy, V. y Zambelli, S. (1991): *Business Cycles: Theories, Evidence and Analysis.*, IEA Conference, Vol. **97**

Tong, H. (1987): “Nonlinear time series models of regularly sampled data: A survey”. *Proc. 1st World Congress of the Bernoulli Society*, **2**, VNU Science Press.

Tong, H. (1990): *Non-linear Time Series. A Dynamical System Approach*, Clarendon Press Oxford.

Wecker, W.E. (1981): “Asymmetric time series”. *Journal of the American Statistical Association*, Vol. **76**, N 373, 16-21.

Capítulo 2

Modelos ARTIMA. El Caso Observable.

Resumen: *En este capítulo introducimos los modelos Autoregresivos Integrados con Threshold en la Media Móvil, que denominaremos modelos ARTIMA. En función de las propiedades de la variable threshold distinguiremos dos clases fundamentales de modelos ARTIMA. Cuando la variable threshold sea observable, hablaremos de modelos ARTIMA-Observables. Cuando la variable threshold sea la perturbación de la serie, y por tanto no observable, hablaremos de modelos ARTIMA-Shock. Para el primer grupo de modelos, ARTIMA-Observables, se estudiarán en detalle las propiedades relacionadas con la persistencia e invertibilidad. Se propondrá un método de estimación para el que se obtienen las condiciones suficientes de consistencia y normalidad asintótica. Igualmente, se presenta la estrategia para el contraste de las hipótesis fundamentales asociadas al modelo. Finalmente, el capítulo se concluye con una aplicación práctica del modelo a la tasa de desempleo de Estados Unidos.*

2.1. Introducción

Es comúnmente aceptado que las series temporales en economía pueden tener respuestas asimétricas a las perturbaciones que forman la serie. Este es el caso del PNB, la tasa de desempleo e inflación, donde la respuesta asimétrica se ha relacionado con el Business Cycle o el signo de la perturbación, ver Breadry y Koop (1994), Falk (1986), Hamilton (1989), Elwood (1998) y Neftci (1984) entre otros. Sin embargo, las asimetrías en la respuesta a las perturbaciones también aparecen en otras series, como determinados precios industriales o stocks y tipos de cambio o interés, ver por ejemplo Wecker (1981), Engle y Smith (1999), Obstfeld y Taylor (1997) y Bec, Salem y Carrasco (2004), entre otros. Estas asimetrías necesitan de modelos no lineales, entre los cuales cabe destacar, por su amplia aplicación, los modelos threshold. Como ya se comentó en la introducción, los modelos threshold se caracterizan por considerar distintos regímenes en la dinámica de la serie, de forma que en cada uno de los regímenes la dinámica es lineal, y el régimen se determina a través de una variable (que denominaremos variable threshold) y un vector de umbrales o thresholds.

A pesar del interés creciente de las asimetrías en las series temporales y la potencial aplicación de los modelos threshold en este contexto, el desarrollo de la teoría asintótica en la estimación e inferencia para determinados casos de los modelos threshold es limitada. Los modelos threshold fueron introducidos por Tong (1983) y Wecker (1981) para el caso particular de medias móviles. Para el caso de threshold en la parte autoregresiva (modelos TAR), algunos resultados sobre consistencia e inferencia asintótica se pueden encontrar en Chan (1993), Hansen (2000), Caner y Hansen (2001) y Gonzalez y Gonzalo (1998) entre otros. Para el caso de threshold en la media móvil, (modelos TMA), hay varios trabajos describiendo estos modelos, ver por ejemplo, Elwood (1998), Gooijer (1998) y Brännäs y Gooijer (1994), aunque no se presentan resultados teóricos sobre las propiedades asintóticas de los estimadores. Solo en el contexto de inferencia indirecta se presentan resultados de este tipo, ver por ejemplo Guay y Scaillet (2003).

Otros modelos no lineales relacionados con los anteriores y utilizados en el estudio de asimetrías, son los modelos de transición suave, caracterizados porque la transición de un régimen a otro se realiza a través de una función suave, normalmente diferenciable. Algunas propiedades y resultados teóricos sobre estos modelos se pueden encontrar, por ejemplo, en Luukkonen, et al. (1988), Granger y Terasvirta (1993), Terasvirta (1994) y Engle y Smith (1999). Finalmente, otra aproximación diferente es la propuesta por Hamilton (1989) en la que el cambio de régimen es el resultado de un proceso de Markov de estados discreto. En este caso, la variable threshold no es observable, pero además no es identificable, con lo que no es posible explicar porque las asimetrías se producen cuando éstas se producen.

Así, el objetivo de este capítulo es introducir una nueva clase de modelos que permita extender el estudio de las asimetrías en la respuesta a las perturbaciones de la serie, incluyendo el caso de respuestas transitorias y permanentes, caso no estudiado hasta el momento. A esta nueva clase de modelos la denominaremos modelos autoregresivos integrados con threshold en la media móvil, ARTIMA. Los modelos ARTIMA poseen un componente de paseo aleatorio en la parte autoregresiva y diferentes regímenes threshold en la parte de medias móviles. En principio, dado que la serie es integrada de orden uno, todas las perturbaciones serán permanentes, sin embargo, permitiendo una raíz unitaria en algún régimen de la media móvil, se puede probar que bajo determinadas condiciones las perturbaciones pertenecientes a ese régimen serán transitorias. Por lo tanto, estos modelos permitirán proponer un nuevo método de identificación de las perturbaciones basado en su comportamiento a largo plazo, permanente o transitorio, y en criterios económicos generales, con la ventaja de que estos criterios son contrastables. Estos criterios se establecen en función de la variable threshold, distinguiendo dos casos fundamentales, cuando la variable threshold es observable, ARTIMA-Observable, y cuando es la propia perturbación, ARTIMA-Shock, y por lo tanto no observable.

En este capítulo se estudiarán las propiedades de invertibilidad y persistencia del mode-

lo ARTIMA-Observable. Igualmente, se prueba que los parámetros del modelo ARTIMA-Observable están identificados, proponiéndose el método de mínimos cuadrados condicionados (CLS) para su estimación. Además, se dan las condiciones para consistencia y normalidad asintótica de dichos estimadores que permitirán la inferencia sobre las hipótesis básicas del modelo. Por lo comentado anteriormente, estas hipótesis básicas serán dos, la existencia de asimetrías y la existencia de perturbaciones transitorias. La estrategia consistirá en un contraste en dos etapas. En primer lugar, se contrastará la existencia de asimetrías o efecto threshold. Bajo la hipótesis nula no habrá efecto threshold, por lo que cuando el parámetro threshold sea desconocido, no estará identificado, con lo que el estadístico propuesto será del tipo supremo, ver Davies (1977, 1987). En segundo lugar, cuando se haya rechazado la nula de ausencia de efecto threshold y los regímenes en la MA no invertibles sean factibles, la existencia de perturbaciones transitorias se hará a través de un contraste para la presencia de raíces unitarias en algún régimen de la MA.

El resto del capítulo se organiza como sigue. En la Sección 2 se presentan los modelos ARTIMA. En la Sección 3 se analizan las propiedades de invertibilidad y persistencia del modelo ARTIMA-Observable. En la Sección 4, se estudia la estimación de dichos modelos. En la Sección 5, se discute la inferencia básica del modelo. La Sección 6 incluye la aplicación del modelo a la tasa de desempleo. La Sección 7 presenta las conclusiones. Finalmente, el Anexo 1 incluye una breve descripción de los diferentes conceptos de dependencia así como determinadas propiedades y resultados de los mismos que se utilizarán en diferentes pruebas a lo largo de la Tesis. Todas las pruebas se encuentran en el Anexo 2, al final del capítulo.

2.2. Modelos ARTIMA.

En su forma general, los modelos autoregresivos integrados con threshold en la media móvil, ARTIMA($p, m, d, 1$), son:

$$\Phi_p(L)(1-L)^d y_t = \mu + \sum_{j=1}^{m+1} ((1-\theta_j L)\varepsilon_t) 1(r_{j-1} < z_{t-1} \leq r_j), \quad (2.1)$$

donde $\Phi_p(L)$ es un polinomio de retardos de orden p con todas sus raíces fuera del círculo unidad, $1(\cdot)$ denota la función indicador, $d \geq 1$ y z_{t-1} es la variable threshold que determina el cambio de régimen. El término de error aleatorio ε_t es una secuencia real *iid* con $E|\varepsilon_t|^{2\gamma}$ para algún $\gamma > 2$. Los parámetros threshold denotados (r_1, \dots, r_m) con $r_0 = -\infty$, $r_{m+1} = \infty$ son tales que $r_i \in \mathfrak{R}_m \forall i = 1, \dots, m$ con $\mathfrak{R}_m = \{(r_1, \dots, r_m) : -\infty < \underline{r} < r_1 < \dots < r_m < \bar{r} < \infty\}$. De esta manera todos los parámetros threshold deben pertenecer al intervalo acotado $[\underline{r}, \bar{r}]$ correspondiente al espacio muestral de la variable threshold. El parámetro d representa el orden de integración de la serie. En principio, los modelos ARTIMA presentan un componente de medias móviles siempre de orden 1. La función de la MA en los procesos ARTIMA es recoger la posibilidad de perturbaciones transitorias en el marco de un proceso integrado. Para ello, es suficiente una MA de orden 1 y $d = 1$. El comportamiento de la serie en el corto plazo se recogerá en el componente autoregresivo. Finalmente, la selección de z_{t-1} como variable threshold en vez de z_t permite que, como veremos más adelante, la persistencia de ε_t dependa de z_t .

A lo largo de la Tesis distinguiremos dos clases fundamentales de modelos ARTIMA. En el primer caso, la variable threshold es función de una variable observable en t , $z_{t-1} = g(z_{t-1}^*)$. Este caso lo denotaremos como ARTIMA-Observable y un ejemplo sería el incremento de la serie, $z_{t-1}^* = (1-L)y_{t-1}$. El segundo caso aparece cuando la variable threshold es una función de la propia perturbación, $z_{t-1} = g(\varepsilon_{t-1})$, y por lo tanto no observable, y que denotaremos por ARTIMA-Shock. Además, en este caso distinguiremos dos tipos de funciones $g(\cdot)$, la función identidad, $g(\varepsilon_t) = \varepsilon_t$, que cuando $m = 1$ y $r_1 = 0$ esta asociada al signo de la variable threshold, y la función valor absoluto, $g(\varepsilon_t) = |\varepsilon_t|$, asociada al tamaño de la perturbación.

Para presentar los resultados principales de los modelos ARTIMA usaremos una versión

más simple del modelo 2.1, donde $m = d = 1$,

$$\Phi_p(L)(1 - L)y_t = \mu + \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & \text{if } z_{t-1} > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & \text{if } z_{t-1} \leq r, \end{cases} \quad (2.2)$$

o de forma más compacta,

$$\Phi_p(L)(1 - L)y_t = \mu + \varepsilon_t - \theta_{t-1} \varepsilon_{t-1}, \quad (2.3)$$

con $\theta_{t-1} = \theta_1$ si $z_{t-1} > r$ e igual a θ_2 en otro caso. Como veremos más adelante, este modelo permitirá, bajo determinadas condiciones que se establecerán en cada caso, la existencia de perturbaciones permanentes y transitorias dependiendo del régimen de la media móvil en el que se generen. En concreto, la existencia de una raíz unitaria en uno de los regímenes de la media móvil será una condición necesaria para este tipo de asimetría.

En general, los threshold en la parte autoregresiva no generarán este tipo de asimetrías. Los threshold en el AR generan asimetrías en la magnitud de la persistencia en el corto plazo, pero a largo plazo, todas las perturbaciones serán transitorias o permanentes, dependiendo de las propiedades de memoria de z_t . En el caso de los threshold en la MA, la posibilidad de una raíz unitaria en uno de los regímenes podrá generar, al cancelar la raíz unitaria de la parte AR, perturbaciones transitorias. Para ello es necesario que la serie de interés, y_t , sea integrada.

En el resto del capítulo estudiaremos las propiedades del modelo ARTIMA-Observable, el método de estimación y las diferentes estrategias de contraste de las hipótesis de interés.

2.3. ARTIMA-Observable

Como se estableció en la sección anterior, a continuación realizamos un análisis detallado de la principales propiedades del modelo ARTIMA-Observable, que corresponde a

la expresión (2.2) con z_t observable. De ahora en adelante requerimos los siguientes supuestos, donde $f(\cdot)$ es la función de densidad de ε_t , $\mathfrak{I}_{t-1} = \sigma(\varepsilon_{t-1}, z_{t-1}, \varepsilon_{t-2}, z_{t-2}, \dots)$ es el σ -algebra generado por las v.a (ε, z) hasta $t-1$ (ver Anexo 1), y $\theta^0 = (\theta_1^0, \theta_2^0, r^0)$ y σ_ε^2 representan los verdaderos valores de los parámetros del proceso.

G.0 ε_t iid $(0, \sigma_\varepsilon)$, con una función de densidad continua $0 < f(\varepsilon_t) < \infty$ para $\forall \varepsilon_t$ y $E|\varepsilon_t|^{2\gamma} < \infty$ con $\gamma > 2$.

G.1 $|\theta_1^0 - \theta_2^0| = \partial^0 > 0$.

G.2 z_t y $x_t = (1-L)y_t$ son estacionarios, ergódicos y ϕ -mixing de tamaño $-a$ con $a > 1$.

G.3 $\Phi_p(L)$ tiene todas las raíces fuera del círculo unidad, y $\mu \in [-\bar{\mu}, \bar{\mu}]$ con $\bar{\mu} < \infty$.

I.0 $\theta_i = 1$ para algún $i = 1, 2$.

I.1 $E[\varepsilon_t | \mathfrak{I}_{t-1}] = 0$ y $E[\varepsilon_t(1 - \theta_t) | \mathfrak{I}_{t-1}] = 0$ para todo t .

A.0 $E(\log(\theta_t^2)) < 0$.

A.1 $\theta^0 \in \Theta_\theta = \{[-1 + \delta, 1 + \delta] \times [-1 + \delta, 1 - \delta] \times [\underline{r}, \bar{r}]\}$, con $\delta > 0$, tal que

$$E\left(\prod_{i=0}^{\bar{q}-1} |\bar{\theta}_{t-i}|^{2\gamma} \middle/ \mathfrak{I}_{t-\bar{q}}\right) < \bar{\lambda} < 1, \quad \text{con} \quad \bar{\theta}_t = (1 + \delta)1(z_t > \underline{r}) + (1 - \delta)1(z_t \leq \underline{r})$$

para algún $\bar{q} < \infty$.

A.2 $E(|\varepsilon_t|^{2\gamma} 1(z' \leq z_t \leq z) / \mathfrak{I}_{t-1}) \leq |z - z'| \bar{\sigma}_{\varepsilon/z}^{2\gamma}$ con $\bar{\sigma}_{\varepsilon/z}^{2\gamma} < \infty$ y $\gamma \geq 1$,

$E(1(z' \leq z_t \leq z) / \mathfrak{I}_{t-1}) \leq |z - z'|M$, con $M < \infty$, y $z' < z \in [\underline{r}, \bar{r}]$ dos valores diferentes de z_t .

A.3 $\min_{z', z} E(|\varepsilon_{t-1}|^{2\gamma} 1(z' < z_{t-1} < z)) \geq |z - z'| \underline{\sigma}_{\varepsilon/z}^2 > 0$

$$M(\underline{r}) = E[1(z_{t-1} < \underline{r}) X_{t-1} X'_{t-1}] > 0$$

$$M(\bar{r}) = E [1(z_{t-1} > \bar{r})X_{t-1}X'_{t-1}] > 0$$

con $X'_{t-1} = [x_{t-1}, \dots, x_{t-p}, 1, \varepsilon_{t-1}]$ y $z' < z \in [\underline{r}, \bar{r}]$ dos valores diferentes de z_t .

La definición de ϕ -mixing se encuentra en el Anexo 1. Los supuestos generales G.0 a G.2 son usuales en los modelos threshold. El supuesto G.3 define el espacio paramétrico de la parte del modelo no relacionada con el componente threshold. I.0 e I.1 son condiciones suficientes para obtener perturbaciones transitorias y permanentes definidas en función de la GI. I.0 implica que, en este sentido, es necesario una raíz unitaria en alguno de los regímenes de la MA, e I.1, que las perturbaciones han de ser una secuencia de martingalas en diferencias, (*mds* en adelante) en el modelo y en cada uno de los regímenes. A.0 establece las condiciones suficientes para la invertibilidad del modelo. Finalmente, A.1 a A.3 se usan en la prueba de consistencia y normalidad asintótica de los estimadores. A.1 describe el espacio paramétrico relacionado con el componente threshold. Nótese que solo el régimen $z_t > r$ puede ser no invertible. Sin embargo, esta condición no excluye la otra posibilidad, dado que $\{z_t < r\} \equiv \{-z_t > -r\}$. A.2 es una cota de los momentos condicionales de ε_t y z_t , y A.3 implica una condición de rango completo que se usa en la identificación de los parámetros, y que fundamentalmente establece la discontinuidad del modelo. Los supuestos A.2 y A.3 están en el espíritu de los supuestos 1.4, 1.5, 1.7 y 1.8 de Hansen (2000).

En las Figuras 2.1 a 2.3, se encuentran representadas varias series correspondientes a modelos ARTIMA-Observables, incluyendo en cada caso un paseo aleatorio. En los gráficos del lado izquierdo se muestran las series para distintos valores de r , θ_i y μ , y en el derecho sus funciones de autocorrelación. Concretamente, los modelos son:

Figura 2.1: $(1-L)y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1}1(|x_{t-1}| > r) - \theta_2\varepsilon_{t-1}1(|x_{t-1}| \leq r)$ con $x_t = (1-L)y_t$.

Figura 2.2: $(1-L)y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1}1(x_{t-1} > r) - \theta_2\varepsilon_{t-1}1(x_{t-1} \leq r)$

Figura 2.3: $(1-L)y_t = \mu + \varepsilon_t - \theta_1\varepsilon_t1(|z_{t-1}| > r) - \theta_2\varepsilon_t1(|z_{t-1}| \leq r)$ con $z_t \sim iidN(0, 1)$.

Los valores que pueden tomar θ_1 y θ_2 son 0 y 1. Si $\theta_i = 1$ denotaremos al régimen i como transitorio, y si $\theta_i = 0$ como permanente. En el gráfico que representa la autocorrelación se muestra el porcentaje de perturbaciones en cada uno de los regímenes. Un hecho destacable es que la función de autocorrelación del paseo aleatorio no muestra siempre el decaimiento más lento, a pesar de ser la serie con mayor porcentaje de perturbaciones permanentes. La razón se encuentra en la varianza estimada de la serie, con la que se estandariza la función de autocorrelación. El decaimiento en la autocorrelación será más lento en la medida en que las perturbaciones permanentes contribuyan menos a la varianza de la serie. Este efecto se puede ver claramente en la Figura 2.1. Otro hecho destacable, es que esta diferencia desaparece cuando introducimos deriva en el modelo.

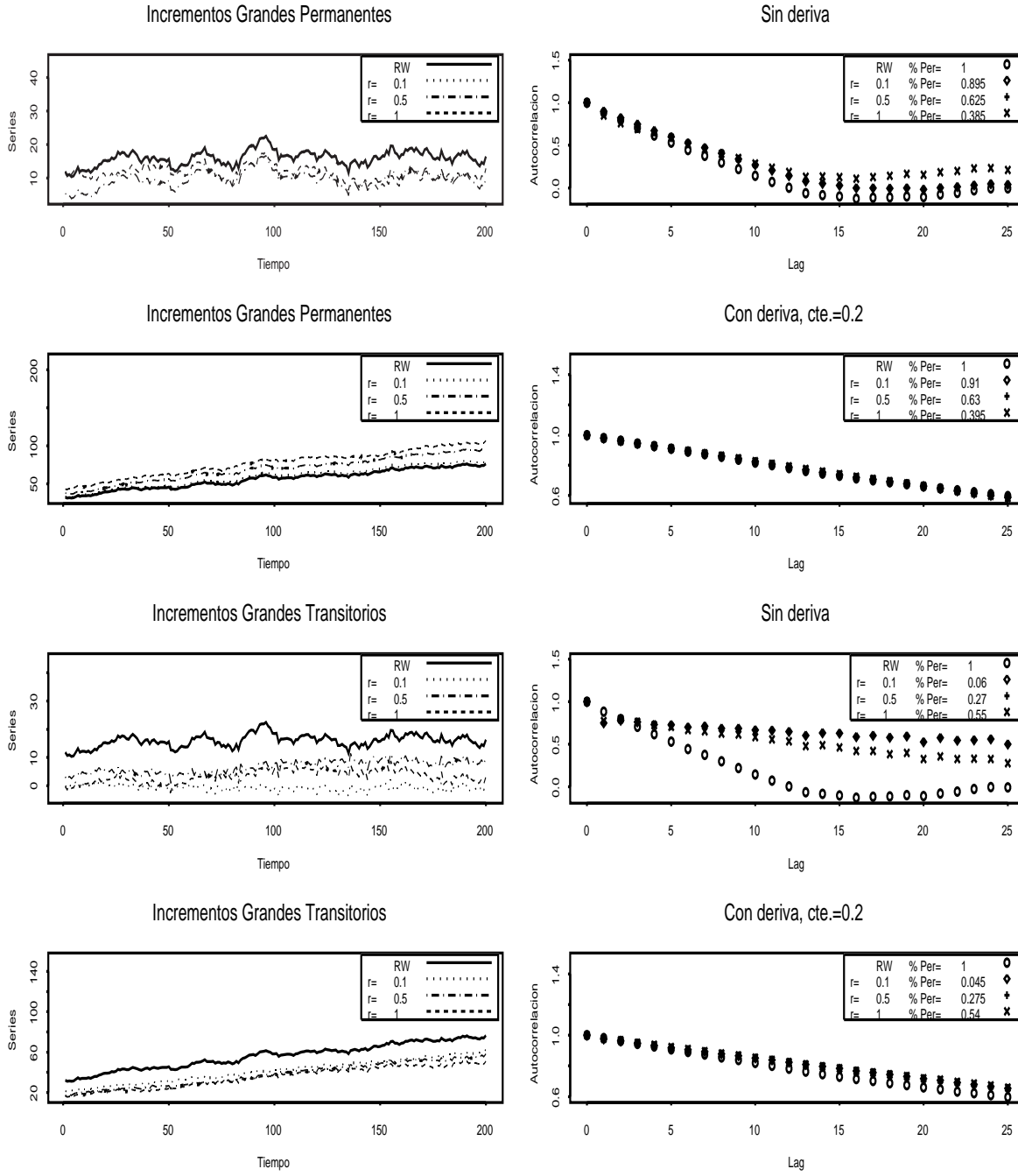
Después de describir las condiciones suficientes que usaremos en el análisis de los modelos ARTIMA-Observables y algunos gráficos de las series, en el resto de esta sección nos centramos en el estudio de la persistencia, a través de la función respuesta a un impulso, y la invertibilidad.

2.3.1. Función Respuesta a un Impulso

Como se comentó en el capítulo introductorio, para definir la persistencia de las perturbaciones, y por tanto, para establecer si una perturbación es permanente o transitoria, nos basaremos en la función respuesta a un impulso general (GI) propuesta por Koop et al. (1996) y Potter (2000). Recordamos la expresión de la GI para una determinada perturbación ε_t e historia w_{t-1} ,

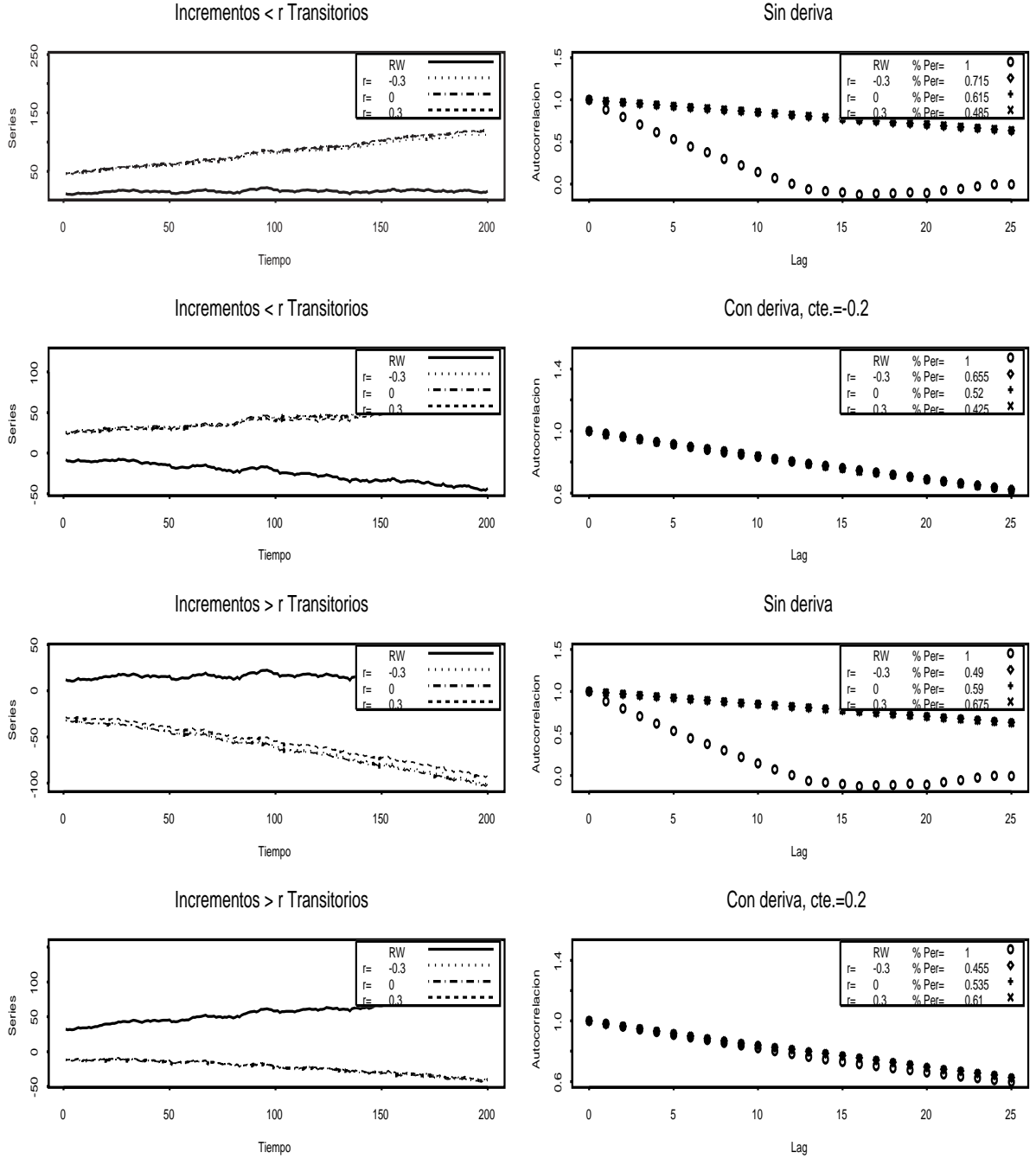
$$GI(k, \varepsilon_t, w_{t-1}) = E[y_{t+k} | \varepsilon_t, w_{t-1}] - E[y_{t+k} | w_{t-1}] \quad \text{para } k = 0, 1, 2, \dots \quad (2.4)$$

Para el caso de los modelos ARTIMA-Observables, es fácil comprobar que la persistencia de cada perturbación dependerá también de z_t . Como z_t es observable en t , proponemos

Figura 2.1: Serie ARTIMA con $z_t = |x_t|$ 

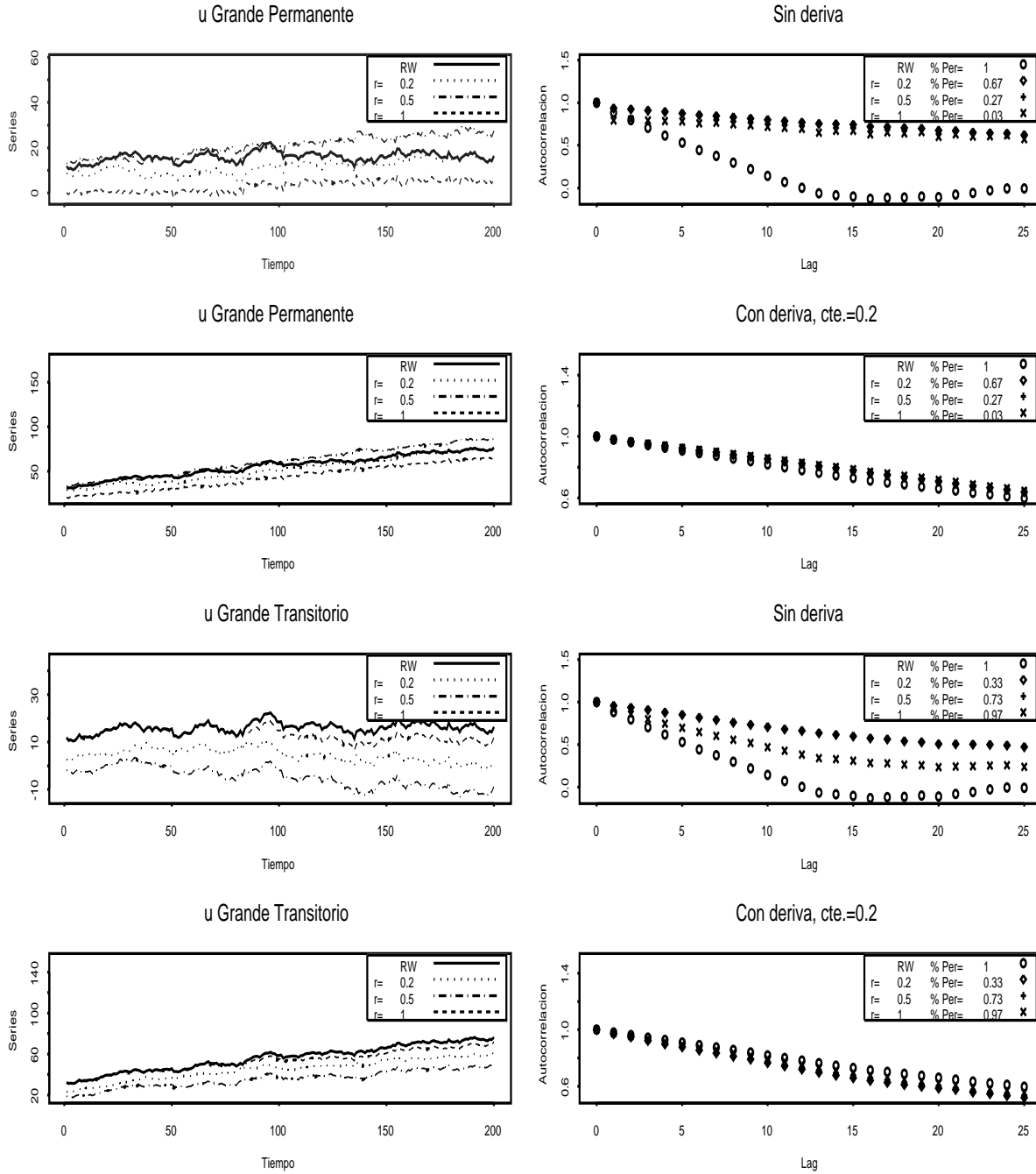
Nota: ARTIMA: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|(1 - L)y_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|(1 - L)y_{t-1}| < r)$; RW: $(1 - L)y_t = \mu + \varepsilon_t$.

Régimen Transitorio: $\theta_i = 1$; Régimen Permanente: $\theta_i = 0$.

Figura 2.2: Serie ARTIMA con $z_t = x_t$ 

Nota: ARTIMA: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1((1 - L)y_{t-1} > r) - \theta_2 \varepsilon_{t-1} 1((1 - L)y_{t-1} < r)$; RW: $(1 - L)y_t = \mu + \varepsilon_t$.

Régimen Transitorio: $\theta_i = 1$; Régimen Permanente: $\theta_i = 0$.

Figura 2.3: Serie ARTIMA con $z_t = |u_t|$ 

Nota: ARTIMA: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|u_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|u_{t-1}| < r)$, $u_t \sim iidN(0, 1)$; RW: $(1 - L)y_t = \mu + \varepsilon_t$.

Régimen Transitorio: $\theta_i = 1$; Régimen Permanente: $\theta_i = 0$.

extender la definición anterior de GI para $v_t = (\varepsilon_t, z_t)$:

$$GI(k, v_t, w_{t-1}) = E[y_{t+k} | v_t, w_{t-1}] - E[y_{t+k} | w_{t-1}] \quad \text{para } k = 0, 1, 2, \dots$$

donde en este caso, w_{t-1} es la historia pasada de v_t . Cuando z_t no sea contemporánea a ε_t , la generalización no será necesaria ya que z_t se podrá incluir en w_{t-1} . Cuando z_t sea contemporánea a la perturbación, no podrá entrar en w_{t-1} , y la generalización será necesaria para una correcta interpretación de la GI.

El siguiente lema establece la relación entre la raíz unitaria en uno de los regímenes de la MA y la existencia de perturbaciones transitorias en los modelos ARTIMA.

Lema 2.1 *Sea y_t un proceso ARTIMA como (2.2) con*

$$\Psi(L) = \sum_{j=0}^{\infty} \psi_j L^j,$$

el polinomio de retardos en su representación $MA(\infty)$ de $\Phi_p(L)x_t = \varepsilon_t$, bajo G.3, I.0 e I.1, la función GI de y_t viene dada por

$$GI(k, v_t, w_{t-1}) = \begin{cases} \psi_{k-1} \varepsilon_t & \text{si } \theta_t = 1 \\ \left[(1 - \theta) \sum_{j=0}^{k-2} \psi_j + \psi_{k-1} \right] \varepsilon_t & \text{si } \theta_t = \theta \neq 1, \end{cases}$$

donde $\psi_0 = 1$ y $\lim_{k \rightarrow \infty} \psi_k = 0$;

Claramente cuando $\theta_t = 1$, lo que implica un régimen MA no invertible, la perturbación ε_t es transitoria, dado que $\lim_{k \rightarrow \infty} \psi_{k-1} \varepsilon_t = 0$. En el otro caso, cuando $\theta_t \neq 1$, la perturbación ε_t será permanente ya que $(1 - \theta) \sum_{j=0}^{k-2} \psi_j$ será diferente de 0 cuando $k \rightarrow \infty$.

Conviene destacar dos comentarios respecto al resultado del lema 2.1. El primero, es que la persistencia de ε_t dependerá de z_t , y no de z_{t-1} . Y el segundo, que bajo las condiciones del lema se obtiene la expresión exacta de la GI, y no es necesaria su estimación. En los modelos no lineales este no será el resultado habitual. La clave para este resultado es el

supuesto I.1. Una condición suficiente para I.1 es que $1(z_t > r)$ no este causado en media en el sentido de Granger por ε_t , esto es, $E[1(z_t \leq r) | \varepsilon_t, \mathfrak{S}_{t-1}] = E[1(z_t \leq r) | \mathfrak{S}_{t-1}]$.

Para un mejor entendimiento del comportamiento de la GI y de sus peculiaridades en el caso de los modelos ARTIMA-Observable, a continuación consideramos varios ejemplos de modelos lineales y no lineales para los que se calcula la GI teórica y estimada.

Ejemplo 1: En este ejemplo consideramos el modelo lineal general,

$$y_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}.$$

Como vimos en el Capítulo 1, la GI para este modelo será,

$$GI(k, \varepsilon_t, w_{t-1}) = \theta_k \varepsilon_t.$$

Por lo tanto, existen dos posibilidades, $\lim_{k \rightarrow \infty} \theta_k = 0$ o $\lim_{k \rightarrow \infty} \theta_k \neq 0$. En el primer caso, todas las perturbaciones son transitorias. En el segundo, todas son permanentes. Las propiedades de persistencia de ε_t sólo dependen de θ_k , ni del pasado ni de sí misma. Para realizar el ejercicio comparativo consideramos el paseo aleatorio:

RW: $y_t = y_{t-1} + \varepsilon_t$,

con la siguiente función GI:

$$GI(k, \varepsilon_t, w_{t-1}) = \varepsilon_t.$$

Claramente, todas las perturbaciones son permanentes.

Ejemplo 2: En este ejemplo consideraremos el siguiente modelo TAR:

$$y_t = \phi_1 y_{t-1} 1(z_{t-1} > r) + \phi_2 y_{t-1} 1(z_{t-1} \leq r) + \varepsilon_t.$$

Como puede probarse fácilmente,

$$\begin{aligned} GI(k, \varepsilon_t, w_{t-1}) &= \sum_{j=0}^{k-1} [E(\phi_{t+k,j} \varepsilon_{t+k-j} | w_{t-1}, \varepsilon_t) - E(\phi_{t+k,j} \varepsilon_{t+k-j} | w_{t-1})] \\ &\quad + E(\phi_{t+k,k} \varepsilon_t | w_{t-1}) - E(\phi_{t+k,k} | \varepsilon_t, w_{t-1}) \varepsilon_t. \end{aligned}$$

con $\phi_t = \phi_1 1(z_t > r) + \phi_2 1(z_t < r)$ y $\phi_{t+k,j} = \prod_{i=1}^j \phi_{t+k-i}$. La función GI en los modelos TAR es no lineal, y puede depender del pasado y de ε_t . En general, será difícil obtener las esperanzas condicionadas que aparecen en la parte derecha de la igualdad, por lo que habrá que estimarlas por simulación. Esta estimación se describe detalladamente en párrafos posteriores. En cualquier caso, bajo la condición I.1,

$$GI(k, \varepsilon_t, w_{t-1}) = E(\phi_t | \varepsilon_t, w_{t-1}) \prod_{j=t+1}^{t+k-1} E(\phi_j | \varepsilon_t, w_{t-1}) \varepsilon_t.$$

Bajo determinadas propiedades de estacionariedad y ergodicidad de z_t la GI convergerá a 0 cuando $k \rightarrow \infty$ si $E(\phi_t) < 1$. Por lo tanto, aunque se obtenga un comportamiento asimétrico en la persistencia, todas la perturbaciones serán transitorias.

Los modelos concretos usados en la comparación son:

TAR 1 : $y_t = y_{t-1} 1(|z_{t-1}| > 0.6) + 0.5 y_{t-1} 1(|z_{t-1}| \leq 0.6) + \varepsilon_t$ con z_t iid $N(0, 1)$ independiente de ε_t .

TAR 2 : $y_t = y_{t-1} 1(|x_{t-1}| > 0.6) + 0.5 y_{t-1} 1(|x_{t-1}| \leq 0.6) + \varepsilon_t$ con $x_t = (1 - L)y_t$.

Ejemplo 3: Finalmente en este ejemplo consideramos los modelos ARTIMA-Observables de la ecuación (2.2) y cuya GI, bajo las condiciones I.0 e I.1, está descrita en el lema 2.1.

Los modelos concretos considerados son:

ARTIMA 1 : $(1 - 0.5L)(1 - L)y_t = \varepsilon_t - 0.5\varepsilon_{t-1} 1(|z_{t-1}| > 0.6) - \varepsilon_{t-1} 1(|z_{t-1}| \leq 0.6)$ con z_t iid $N(0, 1)$ independiente de ε_t .

ARTIMA 2 : $(1 - 0.5L)(1 - L)y_t = \varepsilon_t - 0.5\varepsilon_{t-1} 1(|x_{t-1}| > 0.6) - \varepsilon_{t-1} 1(|x_{t-1}| \leq 0.6)$ con $x_t = (1 - L)y_t$.

En las Figuras 2.4 y 2.6 se representan las estimaciones de la función de densidad de la GI para los modelos RW, TAR 1, ARTIMA 1 y RW, TAR 2, ARTIMA 2 respectivamente.

En todos los casos, ε_t es *iid* $N(0, 1)$. En la estimación de la GI generamos un solo pasado de la serie, de tamaño $T - 1$ con $T = 200$, $B = 500$ replicas de (ε_T^l, z_T^l) , y para cada una de las B replicas de (ε_T^l, z_T^l) , $B1 = 1000$ replicas de $\{(\varepsilon_{T+j}^i, z_{T+j}^i)\}_{j=0}^N$ con $N = 30$. Así, obtenemos $\{GI(k, (\varepsilon_T^l, z_T^l), w_{T-1}), \}_{l=1}^B$ con:

$$GI(k, (\varepsilon_T^l, z_T^l), w_{T-1}) = \hat{E}(y_{T+k}^l / (\varepsilon_T^l, z_T^l), w_{T-1}) - \hat{E}(y_{T+k}^l / w_{T-1})$$

$$\hat{E}(y_{T+k}^l / (\varepsilon_T^l, z_T^l), w_{T-1}) = B1^{-1} \sum_{i=1}^{B1} y_{c,T+k}^{i,l}$$

$$\hat{E}(y_{T+k}^l / w_{T-1}) = B1^{-1} \sum_{i=1}^{B1} y_{T+k}^{i,l}$$

e $y_{c,T+k}^{i,l}$ obtenido de la ecuación del modelo correspondiente, con (ε_T^l, z_T^l) y $\{(\varepsilon_{T+j}^i, z_{T+j}^i)\}_{j=1}^N$, mientras que $y_{T+k}^{i,l}$ se obtiene de la misma ecuación pero con $\{(\varepsilon_{T+j}^i, z_{T+j}^i)\}_{j=0}^N$. Por lo tanto, para estimar la GI de una perturbación perteneciente al régimen $z_t \in A$, usaremos:

$$\hat{E}(y_{T+k}^l / \varepsilon_T^l, z_T^l \in A, w_{T-1}) = B1^{-1} \sum_{i=1}^{B1} y_{c,T+k}^{i,l} 1(z_t^l \in A).$$

En cada ejemplo, calcularemos la GI condicionada a que $|z_t| > r$ y su complementario, $|z_t| < r$. En la Figura 2.4 se muestra la densidad estimada para el primer grupo de modelos, cuando la variable threshold es una v.a independiente, y por lo tanto cumple las condiciones I.0 e I.1. En el caso ARTIMA, podemos ver como la densidad estimada de las perturbaciones para el régimen transitorio, (valores de z_t pequeños), está centrada en cero y su dispersión converge hacia cero. Por el contrario, la dispersión de la densidad estimada de las perturbaciones para el régimen permanente, (valores de z_t grandes), se mantiene constante, replicando el resultado del lema 2.1. Como puede verse en las otras dos columnas de gráficos, correspondientes a los casos del modelo TAR y Paseo Aleatorio, estos modelos no replican este tipo de asimetría. Claramente, el caso del Paseo Aleatorio, por lineal, no presenta asimetría. El caso del modelo TAR, aunque presenta asimetrías en la persistencia en función del tamaño de la variable z_t , las perturbaciones de ambos regímenes son transitorias. Esto se aprecia mejor en la Figura 2.5, donde se representa la esperanza estimada

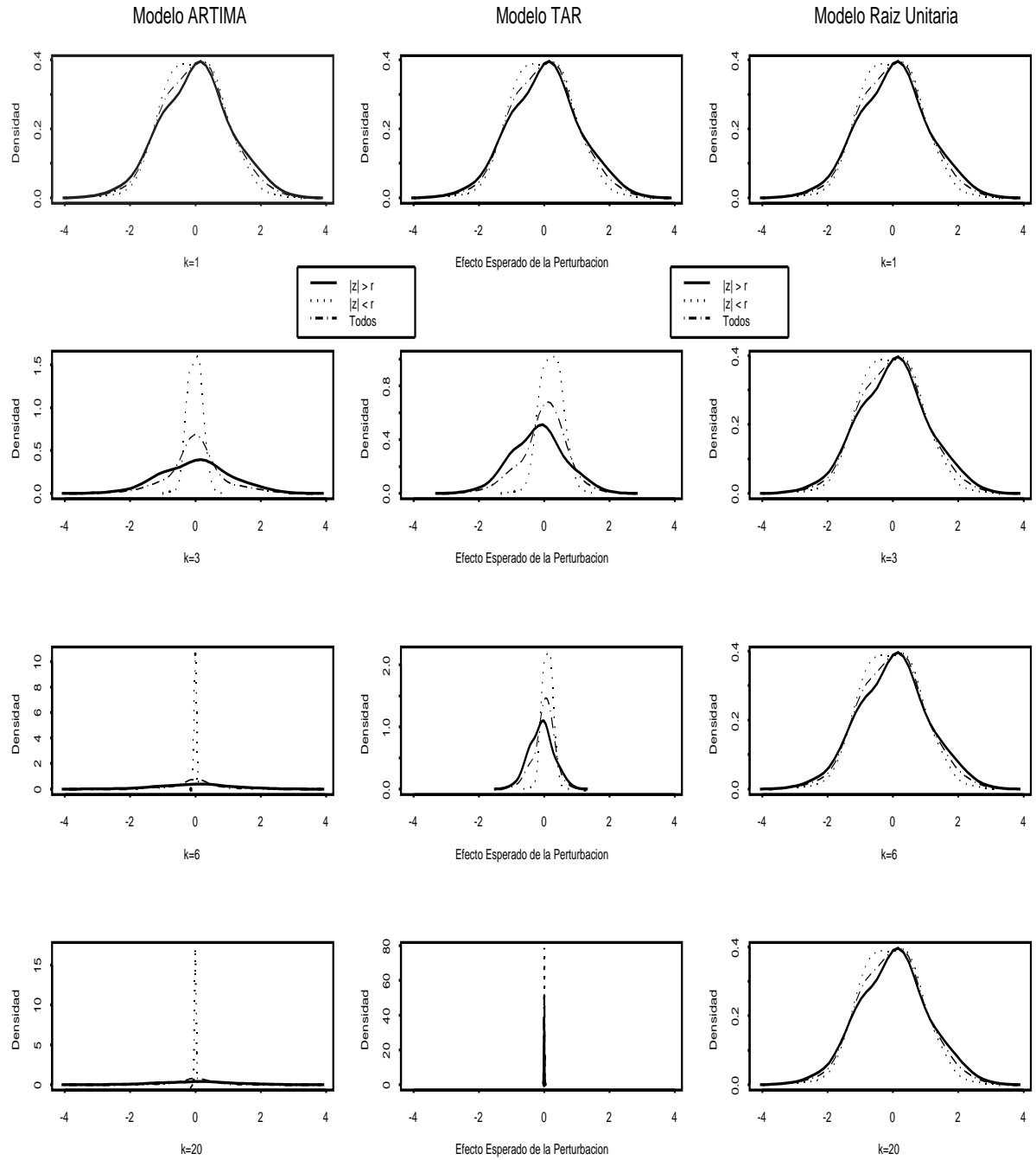
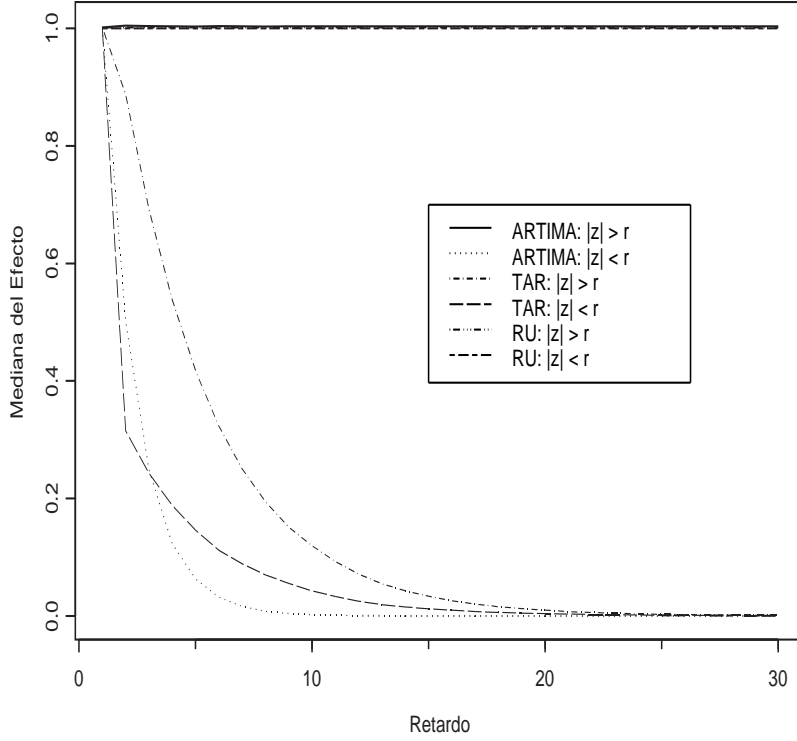
Figura 2.4: Densidad Estimada de la GI. $z_t \sim iid|N(0, 1)|$.

Figura 2.5: Esperanza Estimada de la $\frac{GI}{\varepsilon_t}$. $z_t \sim iid|N(0, 1)|$.



de la GI estandarizada, esto es, $GI(k, v_t, w_t)/\varepsilon_t$. Si no estandarizamos, la esperanza de la GI será cero en cada caso, y no tendrá sentido la comparación. En este gráfico se aprecia claramente como para el modelo TAR, aunque el efecto de las perturbaciones de régimen $|z_t| > 0.6$ decae más lentamente que el resto, también son transitorias, ya que su efecto tiende a cero. Para el modelo ARTIMA, las perturbaciones del régimen $|z_t| > 0.6$ tienen el mismo efecto que en la raíz unitaria, mientras que las del régimen $|z_t| \leq 0.6$ se asemejan a las del modelo TAR, y por lo tanto, son transitorias. La esperanza de la GI se obtiene a través de la mediana muestral para evitar posibles valores extremos al dividir por ε_t .

Para el segundo grupo de modelos, la variable threshold es $|x_t|$, el valor absoluto del incremento de la serie. El resultado es similar al anterior, sin embargo, para el modelo

ARTIMA no se puede concluir que las perturbaciones pertenecientes al régimen de incrementos pequeños, $|x_t| \leq 0.6$, generen efectos transitorios, y esto a pesar de que el régimen es no invertible. La razón de este comportamiento se encuentra en que el modelo ARTIMA con $z_t = |x_t|$ no cumple la condición I.1. Como en el caso anterior, en la Figura 2.7 se representa la esperanza estimada de la GI estandarizada. Claramente, el efecto medio de las perturbaciones del régimen $|x_t| \leq 0.6$ no tiende a cero. Sin embargo, la asimetría en este modelo ARTIMA sigue siendo diferente a la del TAR, ya que la asimetría en la persistencia en el modelo ARTIMA afecta al efecto a largo plazo, $k \rightarrow \infty$, y no solo a corto plazo.

En este estudio, sólo hemos tenido en cuenta las posibles asimetrías provocadas por z_t . Sin embargo, es posible que estos modelos, ARTIMA-Observable y TAR, presenten asimetrías en la persistencia en función de las propiedades de otras variables, como por ejemplo, el signo o tamaño de la perturbación.

2.3.2. Invertibilidad

En esta sección estudiaremos la invertibilidad del modelo ARTIMA-Observable. Como en el caso de la función respuesta a un impulso, la no linealidad elimina la posibilidad de usar la definición clásica de invertibilidad. Por ello, usaremos una definición de invertibilidad más general introducida por Granger y Andersen (1978) y mejorada posteriormente por Hallin (1980). Hallin prueba que bajo no linealidad y con coeficientes constantes, ambas definiciones son equivalentes. Claramente, los modelos ARTIMA entran dentro de esta clase, aunque podrían interpretarse como modelos lineales con coeficientes aleatorios.

La definición de Granger y Andersen es la siguiente,

Definition 2.1 (*Granger y Andersen*) *El proceso $x_t = g(x_{t-1}, \varepsilon_{t-1}, \dots, x_{t-p}, \varepsilon_{t-p}) + \varepsilon_t$ es invertible si*

$$\lim_{t \rightarrow \infty} E(v_t^2) = 0$$

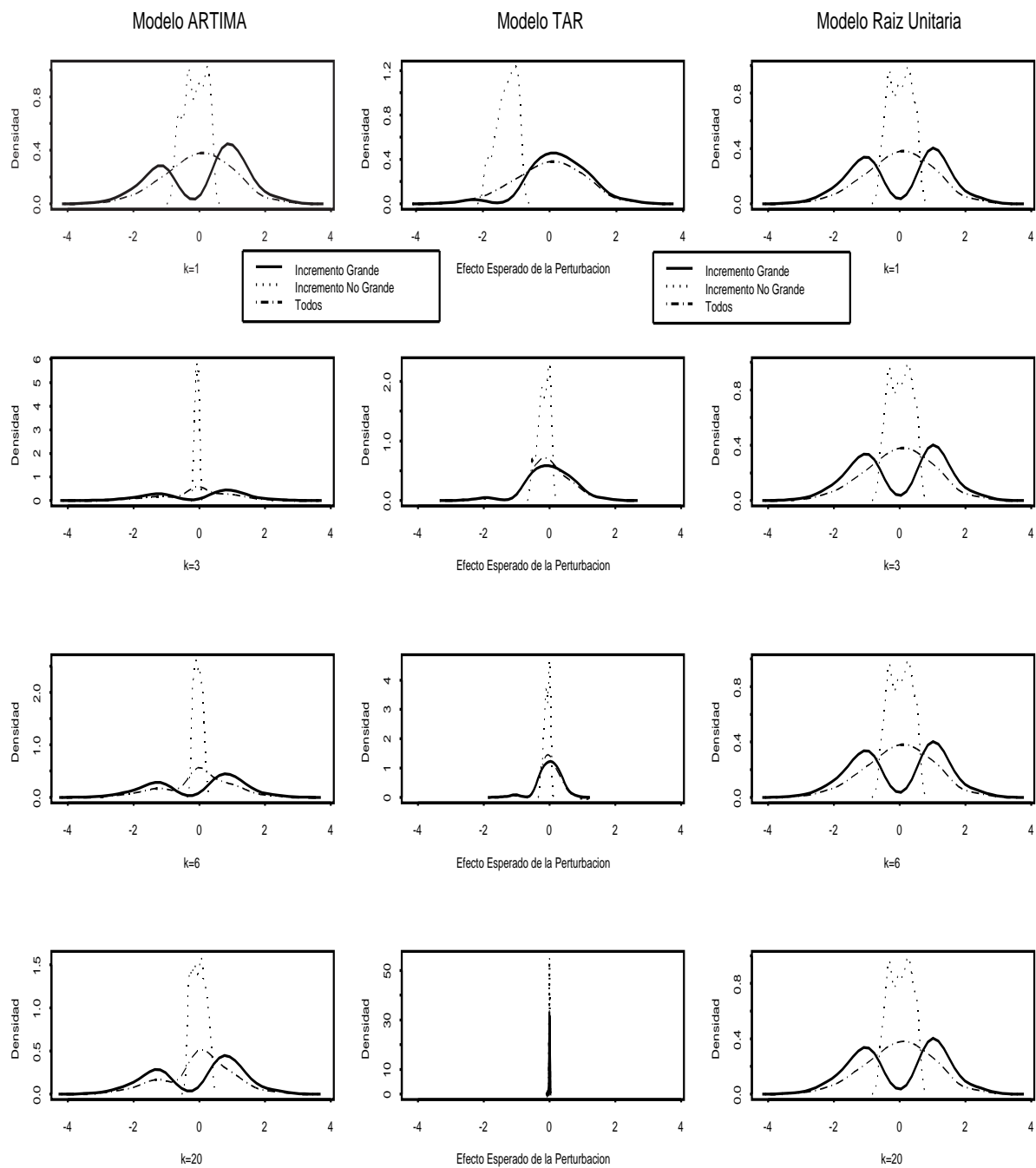
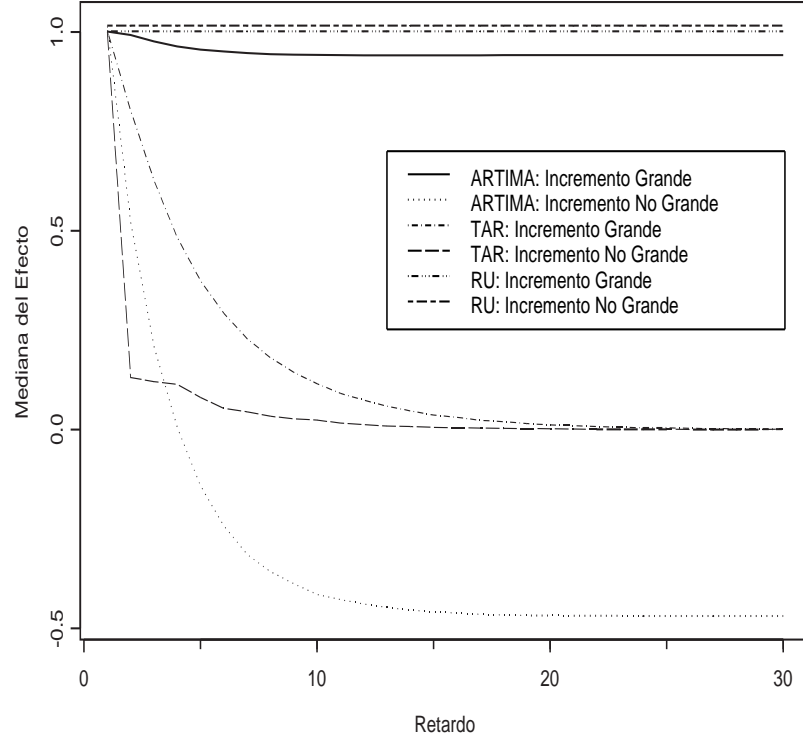
Figura 2.6: Densidad Estimada de la GI. $z_t = |x_t|$.

Figura 2.7: Esperanza Estimada de la $\frac{GI}{\varepsilon_t}$. $z_t = |x_t|$.

con

$$v_t = \varepsilon_t - \hat{\varepsilon}_t = \varepsilon_t - (x_t - g(x_{t-1}, \hat{\varepsilon}_{t-1}, \dots, x_{t-p}, \hat{\varepsilon}_{t-p}))$$

Esta definición de invertibilidad garantiza que en los procesos invertibles las innovaciones se podrán recuperar de la serie observada, de forma que el error cometido tiende a 0 con t . Cuando se trabaja con MA, ya sean lineales o no lineales, y predicción, la invertibilidad se convierte en una propiedad fundamental, ya que las predicciones de la serie dependerán de la innovaciones pasadas, que no son observadas, y por lo tanto, han de obtenerse de la serie observada.

El principal resultado sobre invertibilidad para el modelo ARTIMA-Observable se presenta en el siguiente lema.

Lema 2.2 *Sea y_t un proceso ARTIMA como (2.2), con z_t observable, entonces bajo G.0, G.2 y A.0, y_t es invertible.*

Por lo tanto, a la vista del lema 2.2, es posible que el proceso y_t sea invertible y tener algún régimen MA no invertible. Las condiciones suficientes G.0, G.2 y A.0, son muy similares a las de estacionariedad en los modelos Threshold Autoregresivos, TAR.

2.4. Estimación

En esta sección presentamos el método de estimación propuesto y los resultados de identificabilidad, consistencia y tasa de convergencia de los estimadores. Para el último resultado definimos el orden de convergencia $O_p(T^{-\alpha})$ en el siguiente sentido, X_T es $O_p(T^{-\alpha})$ si para $\forall \varepsilon > 0 \exists \Delta_\varepsilon < \infty$ y $T_\varepsilon \geq 1$ entero tal que $P[|T^\alpha X_T| > \Delta_\varepsilon] < 1 - \varepsilon$ para $\forall T > T_\varepsilon$.

Como método de estimación se propone Mínimos Cuadrados Condicionales (CLS), que es el método propuesto por Chan (1993) para la estimación de los parámetros en los modelos SETAR. Los resultados de interés respecto a los estimadores, se probarán para el siguiente caso particular de modelo ARTIMA-Observable:

$$(1 - \rho L) \Delta y_t = (1 - \rho L) x_t = \mu + \varepsilon_t - \theta_{t-1} \varepsilon_{t-1} = \mu + \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & \text{if } |z_{t-1}| > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & \text{if } |z_{t-1}| \leq r, \end{cases}$$

con $\Delta = (1 - L)$. Las pruebas para el caso mas general, $p > 1$, se siguen directamente a partir de las pruebas desarrolladas en esta Tesis. Así, en este modelo los parámetros a estimar serán cinco, $(\theta_1, \theta_2, \rho, \mu, r)$. En esta sección probaremos que bajo determinadas condiciones y a través de la estimación CLS estos parámetros están identificados y sus estimadores son consistentes.

En general, para los modelos Threshold autoregresivos (TAR), la tasa de convergencia de los estimadores es $T^{1/2}$, excepto para r , el parámetro threshold, que será T . Este parti-

cular resultado para el estimador de r , se debe a la clase de discontinuidad que el modelo presenta en el parámetro r a través de la función indicador.

En el método CLS la función objetivo será la suma de los errores al cuadrado, la cual, definiendo $\beta = (\theta_1, \theta_2, \rho, \mu)$, se puede expresar como,

$$Q_T(\beta, r) = \sum_{t=1}^T e_t^2(\beta, r).$$

con

$$e_t(\beta, r) = x_t - \rho x_{t-1} - \mu + \theta_{t-1} e_{t-1}(\beta, r), \quad e_0 = 0,$$

$$\theta_{t-1} = \begin{cases} \theta_1 & \text{if } z_{t-1} > r \\ \theta_2 & \text{if } z_{t-1} \leq r. \end{cases}$$

Por tanto, el estimador de mínimos cuadrados condicionales (CLSE) de ϕ^0 , con $\phi^0 = (\beta^0, r^0)$ el verdadero valor de los parámetros, se define como $\hat{\phi}_T = (\hat{\beta}_T, \hat{r}_T)$, que debe satisfacer:

$$\hat{\beta}_T(r) = \arg \min_{\beta \in \Theta_\beta} Q_T(\beta, r)$$

y

$$\hat{r}_T = \arg \min_{r \in [-\bar{r}, \bar{r}]} Q_T(\hat{\beta}_T(r), r)$$

donde $\Theta_\beta = \{\beta \text{ t.q. } \theta \in \Theta_\theta, \rho \in [-1 + \delta, 1 - \delta] \text{ y } \mu \in [-\bar{\mu}, \bar{\mu}]\}$ con Θ_θ y δ definidos en el supuesto A.1. Los resultados respecto a $\hat{\phi}_T$ se resumen en el siguiente teorema.

Teorema 2.2 *Bajo G.0-G.3 y A.1-A.3, $\hat{\beta}_{i,T} = \beta_i^0 + O_p(T^{-1/2})$ para $i = 1, \dots, 4$ y $\hat{r}_T = r_0 + O_p(T^{-1})$.*

Prueba, ver Anexo 2.

Este Teorema establece la existencia de un único mínimo en $Q_T(\beta, r)$ alcanzado en ϕ^0 , la consistencia de los CLSE y su tasa de convergencia. Concretamente, los estimadores de los parámetros en los que la función $Q_T(\beta, r)$ es continua, serán consistentes con tasa de

convergencia igual a $T^{1/2}$. Además, el estimador del único parámetro respecto al cual la función objetivo no es continua, r^0 , será superconsistente con tasa T . La discontinuidad de $Q_T(\beta, r)$ en r se establece por el supuesto A.3. Este resultado coincide con el obtenido por Chan (1993) para los modelos SETAR.

2.5. Inferencia

En esta sección estudiamos los resultados necesarios para desarrollar la inferencia relativa al modelo ARTIMA-Observable. De las propiedades de este tipo de modelos, se deduce que dos son las hipótesis principales a contrastar, la existencia de asimetrías en la persistencia de las perturbaciones, y la existencia de perturbaciones transitorias y permanentes. La estrategia de contraste de ambas hipótesis será la siguiente: en primer lugar, contrastaremos la existencia de asimetrías o efecto threshold, donde la hipótesis nula será la no existencia de efecto threshold, esto es, el modelo es lineal. En caso de rechazar esta hipótesis nula, pasamos a contrastar la existencia de perturbaciones transitorias, que como vimos en la Sección 2.3.1, se traduce en el contraste de raíz unitaria en algún régimen de la media móvil, asumido que se cumple I.1.

El resto de la sección se distribuye de la siguiente manera. En primer lugar presentamos los resultados de inferencia asintótica cuando r^0 es conocido. En este caso, el resultado respecto a la distribución asintótica de los estimadores será suficiente para el contraste de las dos hipótesis de interés. En segundo lugar, estudiaremos la inferencia cuando r^0 es desconocido. En este caso, el contraste de existencia de threshold necesitará de resultados adicionales, dado que bajo H_0 , r^0 no está identificado.

2.5.1. Inferencia cuando r^0 es conocido

Conocer el valor del parámetro threshold no es el contexto habitual, excepto cuando las asimetrías vengan provocadas por el signo de alguna variable. En este caso, la elección natural es $r^0 = 0$, y podría estar justificado asumir que r^0 es conocido.

Los resultados relevantes para la inferencia asintótica cuando r^0 es conocido se resumen en los siguientes dos Teoremas que presentamos a continuación. El primer Teorema, dado que r^0 es conocido, muestra la distribución asintótica de $\hat{\beta} = (\hat{\theta}_1, \hat{\theta}_2, \hat{\mu}, \hat{\rho})$. El segundo teorema establece la consistencia del estimador de la matriz de varianzas y covarianzas de la distribución asintótica.

Teorema 2.3 *Sea $\hat{\beta}$ el estimador CLS para el modelo ARTIMA descrito en (2.2), con z_t observable, bajo G.0, G.2, G.3 y A.1-A.3,*

$$T^{1/2} (\hat{\beta} - \beta) \xrightarrow{d} N(\vec{0}, \sigma_\varepsilon^2 H^{-1})$$

con,

$$\begin{aligned} H &= E [h_t(\phi^0) h_t'(\phi^0)] \\ h_t &= (h_t^{(1)}, h_t^{(2)}, h_t^{(3)}, h_t^{(4)})' = \frac{\partial e_t(\beta)}{\partial \beta} \Big|_{\beta=\beta_0} \\ h_t^{(1)} &= \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^+, \quad h_t^{(2)} = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^- \\ h_t^{(3)} &= - \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 x_{t-1-j}, \quad h_t^{(4)} = - \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \\ \prod_{k=1}^0 \theta_{t-k}^0 &= 1 \\ \theta_t^0 &= \theta_1^0 1(z_t > 0) + \theta_2^0 1(z_t < 0) \\ \varepsilon_t^+ &= 1(z_t > r^0) \varepsilon_t \quad y \quad \varepsilon_t^- = \varepsilon_t - \varepsilon_t^+. \end{aligned}$$

Prueba, ver Anexo 2.

Como en el caso estándar de los modelos TAR, la distribución asintótica de los estimadores de los parámetros respecto a los cuales el modelo es continuo, es una normal multivariante de media 0, y matriz de varianzas y covarianzas que depende de las derivadas del proceso $e_t(\beta)$ respecto a los parámetros β . Sin embargo, esta matriz de varianzas es desconocida. Por lo tanto, para poder hacer inferencia basada en esta distribución asintótica, será necesario obtener un estimador consistente de la matriz de varianzas y covarianzas. Este estimador y su consistencia, se proponen en el siguiente teorema.

Teorema 2.4 *Bajo los supuestos del Teorema 2.3*

$$\hat{\sigma}_\varepsilon^2 \frac{1}{T} \sum_{t=2}^T \hat{H}_t^{i,j} \xrightarrow{p} \sigma_\varepsilon^2 H^{i,j}$$

con

$$\begin{aligned} \hat{\sigma}_\varepsilon^2 &= T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 \\ \hat{H}_t^{i,j} &= \hat{h}_t^{(i)} \hat{h}_t^{(j)} \\ \hat{h}_t^{(1)} &= \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \hat{\varepsilon}_{t-1-j}^+, \quad \hat{h}_t^{(2)} = \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \hat{\varepsilon}_{t-1-j}^- \\ \hat{h}_t^{(3)} &= - \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 x_{t-1-j}, \quad \hat{h}_t^{(4)} = - \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \\ \hat{\varepsilon}_t &= e_t(\hat{\beta}), \quad \hat{\varepsilon}_t^+ = \hat{\varepsilon}_t 1(z_t > r^0), \quad \hat{\varepsilon}_t^- = \hat{\varepsilon}_t - \hat{\varepsilon}_t^+ \\ \hat{\theta}_t &= \hat{\theta}_1 1(z_t > r^0) + \hat{\theta}_2 1(z_t < r^0). \end{aligned}$$

Prueba, ver Anexo 2.

Con los resultados presentados en estos dos Teoremas, ya es posible realizar la inferencia relevante sobre el modelo cuando r^0 es conocido. En primer lugar, presentamos el estadístico utilizado para contrastar la existencia de asimetrías o efecto threshold. Este contraste, está basado en la siguiente hipótesis nula:

$$H_0 : \theta_1 = \theta_2,$$

bajo la cual, el modelo es simétrico, frente a la alternativa:

$$H_a : \theta_1 \neq \theta_2.$$

Para construir el contraste, utilizamos el siguiente estadístico:

$$W_{0,T} = \frac{T \left(R' \hat{\beta} \right)^2}{\left[\hat{\sigma}_\varepsilon^2 \left(R' \hat{H}^{-1} R \right) \right]}$$

con $R' = (1, -1, 0, 0)$. Bajo la hipótesis nula de no existencia de efecto threshold convergerá a una χ_1^2 . Este resultado sale directamente de la aplicación de los Teoremas 2.3 y 2.4. Así, rechazaremos la H_0 al nivel de significación α cuando $W_{0,T} > \chi_{1,\alpha}$, con $\chi_{1,\alpha}$ tal que $P[\chi_1^2 > \chi_{1,\alpha}] = \alpha$. Conviene hacer notar que en este contraste los supuestos del Teorema han de cumplirse bajo H_0 . La hipótesis nula afecta fundamentalmente a G.2. En el caso en que $z_t = x_t$ y $p = 0$, para que x_t sea $\phi - mixing$ bastará con que lo sea ε_t . Si $p > 0$, entonces serán necesarias condiciones adicionales sobre ε_t , ver Doukham (1995), Makkaden (1990) y Davidson (1994).

En un segundo paso, si rechazamos la hipótesis nula anterior, contrastamos la existencia de perturbaciones transitorias en un determinado régimen. En este caso, la hipótesis nula es:

$$H_0 : \theta_i = 1,$$

frente a la alternativa:

$$H_a : \theta_i \neq 1,$$

para $i = 1, 2$. El estadístico se basará en el siguiente estadístico:

$$W_{1,T} = \frac{T \left(\hat{\theta}_i - 1 \right)^2}{\left[\hat{\sigma}_\varepsilon^2 \left(\hat{H}^{-1}(1, 1) \right) \right]},$$

que bajo H_0 convergerá a una χ_1^2 . Como en el estadístico anterior, este resultado se deriva directamente de la aplicación de los Teoremas 2.3 y 2.4. Rechazaremos la H_0 al nivel de significación α cuando $W_{1,T} > \chi_{1,\alpha}$ con $\chi_{1,\alpha}$ definido anteriormente.

Este contraste se desarrolla bajo dos supuestos implícitos. El primero, es la existencia de una raíz unitaria en el componente AR. El segundo, hace referencia al hecho de que el régimen de perturbaciones transitorias es conocido y se cumple el supuesto I.1 bajo la nula.

2.5.2. Inferencia cuando r^0 es desconocido

Como decíamos en la sección anterior, excepto para el caso de asimetrías por razón de signo, r^0 suele ser desconocido. En esta sección, se obtienen los resultados necesarios para la implementación de los contrastes descritos anteriormente cuando r^0 es desconocido. En primer lugar, se deriva la distribución asintótica de $\hat{\beta}(\hat{r})$ bajo el supuesto de existencia de efecto threshold. Este resultado, que se apoya en la superconsistencia de \hat{r} , permite contrastar la existencia de perturbaciones transitorias. Sin embargo, no será válido para contrastar la hipótesis nula de efecto threshold, ya que bajo esta hipótesis el parámetro r^0 no está identificado, y por tanto, \hat{r} no será consistente.

Así, en el siguiente Teorema se muestra que cuando r^0 es desconocido y existe efecto threshold, la distribución asintótica de $\hat{\beta}$ es independiente de \hat{r} . Este resultado es equivalente al obtenido por Chan (1993) para los modelos SETAR.

Teorema 2.5 *Bajo los supuestos G.0-G.3 y A.1-A.3,*

$$T^{1/2} \left(\hat{\beta}(\hat{r}) - \beta \right) \xrightarrow{d} N \left(\vec{0}, \sigma_\varepsilon^2 H^{-1} \right)$$

donde H se definió en el Teorema 2.3.

Prueba, ver Anexo 2.

Con este resultado, el estadístico para el contraste de existencia de raíz unitaria presentado en la sección anterior sigue siendo válido si sustituimos r^0 por \hat{r} . Igualmente, la matriz de varianzas y covarianzas descrita en el Teorema 2.4, sigue siendo consistente cuando utilizamos \hat{r} . La prueba de consistencia de este estimador se encuentra dentro de la prueba del Teorema 2.5.

Contraste para la existencia de Threshold cuando r^0 es desconocido

A continuación presentamos el estadístico de contraste para la existencia de efecto threshold cuando r^0 es desconocido, así como la teoría asintótica necesaria para su desarrollo. Como comentamos anteriormente, bajo la hipótesis nula de simetría o no threshold, el parámetro r no estará identificado, y los resultados anteriores no serán válidos. Este tipo de contraste se engloba dentro de la clase de contrastes con parámetros ruidosos presentes solo bajo la alternativa. La literatura sobre este tipo de contrastes es bastante amplia, véase, por ejemplo, Andrews (1993), Andrews y Ploberger (1994), Davies (1987) y Hansen (1996) entre otros.

La nueva estrategia de contraste se basará en la siguiente ecuación de regresión modificada:

$$x_t = \rho x_{t-1} + \mu - \theta \varepsilon_{t-1} - \alpha \varepsilon_{t-1} 1(z_{t-1} < r) + \varepsilon_t, \quad (2.5)$$

donde los parámetros de la ecuación, $\beta = (\theta, \alpha, \mu, \rho)$, se estimarán por CLS. La razón fundamental de modificar la ecuación de regresión es que en esta ecuación solo uno de los regresores depende de r , lo que facilita la demostración de los resultados obtenidos. Así, en este caso,

$$e_t(\beta, r) = x_t - \rho x_{t-1} - \mu + \theta e_{t-1}(\beta, r) + \alpha(r) e_{t-1}(\beta, r) 1(z_{t-1} \leq r) \quad \text{con } e_0 = 0.$$

Utilizando esta regresión, la hipótesis de simetría respecto a z_t , es:

$$H_0 : \alpha(r) = 0.$$

Análogamente, la hipótesis alternativa de asimetría será:

$$H_a : \alpha(r) \neq 0.$$

Por lo tanto, un estadístico natural para este contraste será $\Gamma_T(r) = \frac{T\hat{\alpha}^2(r)}{\hat{V}(\hat{\alpha}(r))}$, con:

$$\begin{aligned}\hat{V}(\hat{\alpha}(r)) &= \hat{\sigma}_\varepsilon^2 R' H_T^{-1}(\hat{\beta}, r) R \\ R' &= (0, 1, 0, 0) \\ H_T^{-1}(\hat{\beta}, r) &= T^{-1} \sum_{t=1}^T h_t(\hat{\beta}, r) h_t'(\hat{\beta}, r) \\ h_t(\hat{\beta}, r) &= \left(\sum_{j=1}^t \theta^{0j-1} e_{t-j}(\hat{\beta}, r), \sum_{j=1}^t \theta^{0j-1} e_{t-j}^-(\hat{\beta}, r), -\sum_{j=1}^t \theta^{0j-1} x_{t-1}, -\sum_{j=1}^t \theta^{0j-1} \right) \\ e_{t-j}^-(\hat{\beta}, r) &= 1(z_{t-j} < r) e_{t-j}(\hat{\beta}, r).\end{aligned}$$

Dado que bajo la hipótesis nula r no está identificado, proponemos contrastar H_0 con el $\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r)$, como sugiere Davies (1977, 1987). Para estudiar el comportamiento asintótico de este estadístico definimos previamente los siguientes procesos. En primer lugar, $s(r)$, que será un proceso gaussiano de media cero y función de covarianzas igual a $K(r_1, r_2) = \sigma_\varepsilon^2 E(h_t(\beta^0, r_1) h_t'(\beta^0, r_2))$ con $h_t^{(i)}(\beta^0, r) = \frac{\partial e_t(\beta, r)}{\partial \beta_i} \Big|_{\beta_i = \beta_i^0}$. En segundo lugar, $\bar{S}(r) = R' H^{-1}(r) s(r)$, que de nuevo será un proceso gaussiano de media cero pero cuya función de covarianzas es $\bar{K}(r_1, r_2) = R' H^{-1}(r_1) K(r_1, r_2) H^{-1}(r_2) R$, con $H(r) = E(h_t(\beta^0, r) h_t'(\beta^0, r))$ y $R' = (0, 1, 0, 0)$.

El siguiente Teorema muestra la distribución asintótica del estadístico propuesto en esta sección, donde \Rightarrow implica convergencia uniforme en distribución.

Teorema 2.6 *Bajo H_0 y los supuestos G.0, G.2, G.3, y A.1-A.3, para cualquier funcional continuo $G(\cdot) : [\underline{r}, \bar{r}] \rightarrow \mathbb{R}$, se tiene que*

$$G(\Gamma_T(r)) \Rightarrow G\left(\bar{S}'(r) \bar{K}^{-1}(r, r) \bar{S}(r)\right),$$

Prueba, ver Anexo 2.

Dado que el $\sup_{r \in [\underline{r}, \bar{r}]}$ es un funcional continuo de $[\underline{r}, \bar{r}]$ en \mathbb{R} tenemos que bajo H_0 :

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\hat{\alpha}^2(r)}{\hat{V}(\hat{\alpha}(r))} \Rightarrow \sup_{r \in [\underline{r}, \bar{r}]} \bar{S}'(r) \bar{K}^{-1}(r, r) \bar{S}(r).$$

El Teorema 2.6 muestra la distribución asintótica de $\Gamma_T(r)$, que vendrá dada por $\bar{S}'(r) \bar{K}^{-1}(r, r) \bar{S}(r)$, que para un r fijo, coincidirá con una χ^2 , ver el caso de $r^0 = 0$ conocido. En general, la distribución bajo H_0 de $G(\Gamma_T(r))$ dependerá de la función de covarianzas $\bar{K}(r_1, r_2)$, y los valores críticos no podrán ser tabulados excepto en casos especiales. Por lo tanto, para poder realizar el contraste, proponemos una aproximación bootstrap que nos permita calcular el p -valor del estadístico, y que describimos a continuación.

De $\hat{\varepsilon}_t$, el estimador CLS de ε_t en la ecuación (2.5), generamos M muestras de $\{\eta_t^i\}_{t=1}^T$, $\{v_t^i\}_{t=1}^T$ y $\{x_t^i\}_{t=1}^T$, con η_t^i iid $(0, 1)$, $v_t^i = \eta_t^i \hat{\varepsilon}_t$ y $x_t^i = \hat{\rho} x_{t-1}^i + \hat{\mu} - \hat{\theta} v_{t-1}^i + v_t^i$ para $i = 1, \dots, M$, de forma que las series $\{x_t^i\}_{t=1}^T$ estarán construidas bajo H_0 . Para cada muestra $\{x_t^i\}_{t=1}^T$, estimamos por CLS en la ecuación (2.5), $\tilde{\alpha}^i(r)$ y $\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\tilde{\alpha}^2(r)}{\tilde{V}(\tilde{\alpha}(r))}$. Por lo tanto, tomando la distribución empírica del $\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\tilde{\alpha}^2(r)}{\tilde{V}(\tilde{\alpha}(r))}$ como una aproximación de la distribución del $\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r)$, rechazaremos H_0 si

$$\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r) > q_\epsilon,$$

donde $\epsilon = \hat{P} \left[\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\tilde{\alpha}^2(r)}{\tilde{V}(\tilde{\alpha}(r))} > q_\epsilon \right]$, y \hat{P} es la distribución empírica del $\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\tilde{\alpha}^2(r)}{\tilde{V}(\tilde{\alpha}(r))}$.

Cuando la nula de simetría en z_t sea rechazada, entonces podremos pasar a contrastar la hipótesis de existencia de perturbaciones transitorias, utilizando el estadístico descrito en la sección anterior.

Finalmente, el resultado del Teorema permite extender el análisis a cualquier otro estadístico basado en un funcional continuo de $\Gamma_T(r)$, como es el caso de los propuestos por Andrews y Ploberger (1994) que utilizan distintas medias de $\Gamma_T(r)$.

2.5.3. Comportamiento de los contrastes en muestras finitas

En esta sección se examina, a través de un estudio de Monte Carlo, el comportamiento en muestras finitas de los contrastes para la existencia de threshold propuestos en las secciones anteriores, en concreto, para los casos r^0 conocido y desconocido. El estudio se ha realizado para el caso en que la variable threshold es el incremento de la serie, esto es, para $z_t = g(x_t)$ en la ecuación (2.2) con $x_t = (1 - L)y_t$, $\Phi_p(L) = 1$ y $\mu = 0$. Por motivos computacionales, el tamaño del Monte Carlo seleccionado fue $M = 400$ y $B = 400$ para las replicaciones del Wild Bootstrap. Todos los casos se analizarán para los tamaños muestrales $T = 200$ y $T = 600$ y los niveles de significación $\alpha = 0.05$ y $\alpha = 0.10$. Los cálculos fueron llevados a cabo con el programa informático Digital Visual Fortran 6.

Para estudiar el tamaño y potencia del contraste de existencia de threshold cuando r^0 es conocido, desarrollado en la Sección 2.5.1, nos centraremos en el caso del signo y por lo tanto, $r^0 = 0$. Para el tamaño se ha considerado el proceso ARIMA(0,1,1):

$$(1 - L)y_t = x_t = \varepsilon_t - \theta\varepsilon_{t-1},$$

donde ε_t son innovaciones *iid* $N(0, 1)$ y $\theta = (-0.5, 0, 0.5)$. Dado que para r^0 conocido la distribución asintótica es conocida y tabulable, el contraste se realiza utilizando los valores críticos asintóticos (obtenidos de una χ^2 de 1 grado de libertad) y los obtenidos a través del procedimiento Wild bootstrap descrito en la sección anterior. Los resultados se recogen en la Tabla 2.1. Como se puede comprobar la aproximación Wild Bootstrap es sensiblemente mejor que la aproximación asintótica. Sin embargo, ambos casos presentan distorsiones para los valores de $\theta = (-0.5, 0.5)$, debidos probablemente al tamaño del estudio de Monte Carlo. Para $\theta = -0.5$ el contraste resulta conservador, obteniéndose el resultado opuesto para $\theta = 0.5$.

Para el estudio de la potencia se genera el proceso ARTIMA:

$$(1 - L)y_t = x_t = \varepsilon_t - \theta_1\varepsilon_{t-1}1(x_{t-1} > 0) - \theta_2\varepsilon_{t-1}1(x_{t-1} \leq 0)$$

donde ε_t son innovaciones *iid* $N(0, 1)$, $\theta_1 = 0$ y $\theta_2 = (-0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8)$. Los resultados se muestran en las Tablas 2.3 y 2.4 para el tamaño $T = 200$ y $T = 600$ respectivamente. A parte del porcentaje de rechazos para los dos tamaños del contraste, 0.05 y 0.10, y para ambas aproximaciones de los valores críticos, se muestra el porcentaje medio de incrementos negativos y la varianza media de las innovaciones en dicho regimen para cada uno de los valores de θ_2 . Como cabía esperar, a medida que aumenta $\partial = |\theta_1 - \theta_2|$ aumenta la potencia del contraste. Además, el resultado es simétrico entorno a θ , a pesar de que $P(x_t < 0)$ y $E(\varepsilon_t^2 1(x_t < 0))$ son menores para los valores positivos de θ_2 . Este resultado puede deberse al sesgo conservador del contraste para los valores de $\theta < 0$.

Para el caso en que el parámetro threshold r^0 es desconocido, y por lo tanto, para el estadístico $\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r)$ desarrollado en la Sección 2.5.2, nos centraremos en el caso del tamaño del incremento de la serie, esto es, el proceso de la ecuación (2.2) con $z_t = |x_t|$. En el caso de r^0 desconocido, aunque la distribución asintótica es conocida, no es tabulable, y por lo tanto, usaremos unicamente la aproximación Wild Bootstrap. Para el estudio del tamaño usaremos los procesos ARIMA(0,1,1) descritos anteriormente en el caso de r^0 conocido. Los resultados se muestran en la Tabla 2.2. Esta confirma el buen comportamiento del contraste respecto al tamaño.

Para el estudio de la potencia se genera el proceso ARTIMA,

$$(1 - L)y_t = x_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|x_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|x_{t-1}| < r)$$

donde ε_t son *iid* $N(0, 1)$, $(\theta_1, \theta_2) = (0, 1)$ y $r = (0.35, 0.5, 0.7)$. Los resultados se muestran en la tabla 2.5. Como cabía esperar, fijados (θ_1, θ_2) , a medida que aumenta r la potencia del contraste es mayor.

Cuadro 2.1: Tamaño del contraste de existencia de Threshold, $r^0 = 0$.

θ	Valores Críticos Bootstrap				Valores Críticos Asintóticos			
	T=200		T=600		T=200		T=600	
	$\alpha - level$				$\alpha - level$			
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
-0.5	0.040	0.077	0.027	0.067	0.032	0.090	0.030	0.067
0.0	0.050	0.095	0.042	0.080	0.062	0.105	0.040	0.070
0.5	0.057	0.130	0.062	0.135	0.100	0.167	0.085	0.152

Cuadro 2.2: Tamaño del contraste de existencia de Threshold. r^0 desconocido.

	T=200		T=600	
	$\alpha - level$			
θ	0.05	0.10	0.05	0.10
-0.5	0.042	0.100	0.047	0.095
0.0	0.070	0.165	0.047	0.100
0.5	0.050	0.067	0.045	0.110

2.6. Aplicación a la tasa de desempleo.

La literatura tradicional en macroeconomía suele distinguir entre la tasa de desempleo actual y la de equilibrio, de forma que la última no se ve afectada por la primera. De acuerdo con esto, los movimientos en la tasa de desempleo causados por perturbaciones en la demanda serán movimientos de corto plazo, y en el largo plazo, la tasa de desempleo regresa a su tasa natural o de equilibrio. En esta literatura, la tasa natural de desempleo está determinada por las condiciones institucionales del mercado de trabajo. Sin embargo,

Cuadro 2.3: Potencia del contraste de existencia de Threshold, $r^0 = 0$. $T = 200$

$\theta_1 = 0$	θ_2							
	-0.8	-0.6	-0.4	-0.2	0.2	0.4	0.6	0.8
$\alpha - level$	Valor Crítico Bootstrap							
0.05	1.000	0.885	0.505	0.120	0.110	0.445	0.885	0.995
0.10	1.000	0.930	0.610	0.255	0.210	0.575	0.945	0.985
$\alpha - level$	Valor Crítico Asintótico							
0.05	1.000	0.865	0.460	0.145	0.125	0.425	0.840	1.000
0.10	1.000	0.910	0.580	0.255	0.210	0.555	0.905	1.000
$P(x_t < 0)$	0.594	0.581	0.560	0.531	0.469	0.441	0.420	0.403
$\text{Var}(\varepsilon_t 1(x_t < 0))$	0.445	0.420	0.397	0.382	0.380	0.383	0.384	0.380

Cuadro 2.4: Potencia del contraste de existencia de Threshold, $r^0 = 0$. $T = 600$

$\theta_1 = 0$	θ_2							
	-0.8	-0.6	-0.4	-0.2	0.2	0.4	0.6	0.8
$\alpha - level$	Valor Crítico Bootstrap							
0.05	1.000	1.000	0.940	0.355	0.285	0.880	1.000	1.000
0.10	1.000	1.000	0.965	0.490	0.400	0.955	1.000	1.000
$\alpha - level$	Valor Crítico Asintótico							
0.05	1.000	1.000	0.920	0.360	0.270	0.865	1.000	1.000
0.10	1.000	1.000	0.960	0.475	0.420	0.945	1.000	1.000
$P(x_t < 0)$	0.594	0.580	0.559	0.531	0.469	0.442	0.420	0.405
$\text{Var}(\varepsilon_t 1(x_t < 0))$	0.417	0.392	0.370	0.356	0.354	0.358	0.358	0.355

Cuadro 2.5: Potencia del contraste de existencia de Threshold. r^0 desconocido.

$\theta_1 = 0, \theta_2 = 1$	T=200			T=600		
	r					
	$\alpha - level$	0.35	0.5	0.7	0.35	0.5
0.05	0.195	0.485	0.975	0.540	0.985	1.000
0.10	0.360	0.705	0.995	0.690	1.000	1.000
$P(x_t < r)$	0.268	0.370	0.490	0.271	0.375	0.495
$\text{Var}(\varepsilon_t 1(x_t < r))$	0.014	0.040	0.104	0.014	0.042	0.106

la experiencia europea muestra que esta visión convencional no captura aspectos fundamentales de la realidad. Durante la crisis del petróleo, Europa sufrió un período de crecimiento del desempleo. En los 80's, la tasa de desempleo alcanzo prácticamente el doble de la tasa de equilibrio de los 70's, sin embargo, hay una ausencia de perturbaciones adversas que puedan explicar este incremento. Por lo tanto, se hizo necesario el desarrollo de teorías sobre el desempleo alternativas que fueran capaces de explicar la alta persistencia que presenta la tasa de desempleo. Estas teorías desarrollan la idea de que la tasa de equilibrio del desempleo depende de su historia. Tales teorías se enmarcan en lo que se ha denominado histéresis.

En esta sección nos centraremos en dos explicaciones de la histéresis, debidas a Blanchard y Summers (1986) y Ball (1999), y que se centran en el funcionamiento del mercado de trabajo y la relación entre desempleo y salario. La explicación propuesta por Blanchard y Summers, y que denominaremos la teoría de los “miembros”, esta basada en la distinción entre insiders y outsiders, y explora la idea de que los salarios son establecidos por los insiders más que por los outsiders. La propuesta por Ball, es la teoría de duración, y distin-

que entre desempleo de corto y largo plazo, estableciendo que el desempleo de largo plazo ejerce escasa presión sobre el salario. En ambas teorías, un incremento (o disminución) relativamente grande del desempleo podría afectar a la persistencia del desempleo así como a su tasa natural.

En el modelo de Blanchard y Summers, el salario es establecido únicamente por los miembros de un grupo, que normalmente se identifica con un sindicato. Dependiendo de las relaciones entre los miembros y los desempleados podemos obtener distintos grados de persistencia en la tasa de desempleo. En el caso extremo en el que los miembros sean sólo las personas empleadas, la tasa de desempleo se comportará como un paseo aleatorio con deriva, esto es, cualquier cambio en la tasa de desempleo será permanente. En el caso opuesto, cuando uno siempre forma parte del grupo, la tasa de desempleo será un ruido blanco. En los casos intermedios, cuando son necesarios varios periodos para que un desempleado (empleado) pierda (gane) su estatus de miembro, cambios duraderos y grandes de la tasa de desempleo pueden provocar cambios permanentes en el desempleo.

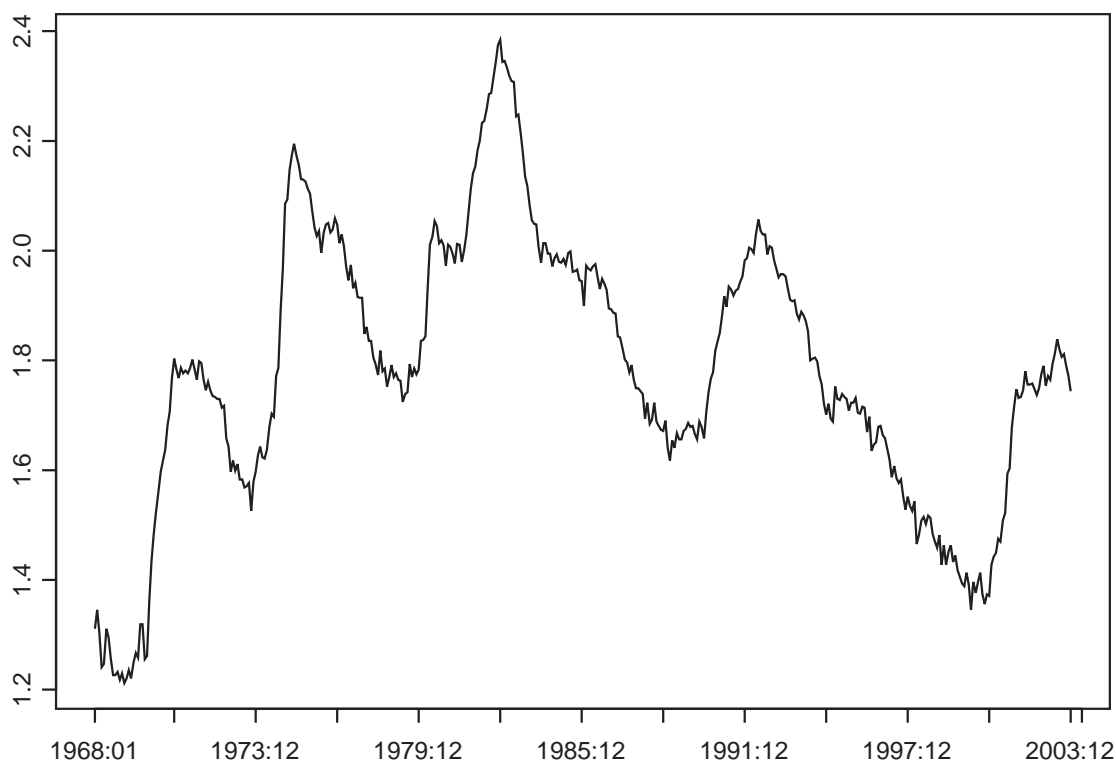
El modelo de duración de Ball distingue entre el desempleo de corto y largo plazo. Este último está caracterizado por pérdidas en el capital humano y de relación con la fuerza de trabajo, de forma que los desempleados de largo plazo se sitúan entre los últimos a contratar, impidiéndoles hacer presión en el salario. Así, las firmas preferirán (o simplemente es más fácil) contratar desempleados de corto plazo. Sin embargo, en los casos en los que haya un elevado nivel de demanda, las empresas preferirán contratar a un desempleado de largo plazo que no contratar. Por lo tanto, la probabilidad de ser vuelto a contratar en un desempleado de corto plazo dependerá del número de desempleados de corto plazo, pero no de los de largo plazo. Claramente, pequeñas variaciones en el volumen de desempleados de corto plazo provocarán variaciones transitorias. Sin embargo, un incremento elevado de la tasa de desempleo implicará un aumento de la probabilidad de pasar a ser un desempleado de largo plazo, lo cual provocaría un cambio permanente en la tasa de desempleo.

Recíprocamente, una fuerte disminución de la tasa de desempleo implica que las empresas contratan desempleados de largo plazo, produciendo una disminución de la tasa natural de desempleo.

En esta sección estudiamos la tasa de desempleo para Estados Unidos (EEUU) y contrastamos la hipótesis de que la persistencia en la tasa de desempleo depende de su incremento, presentando por tanto, histéresis. En concreto contrastaremos, como se deriva de los modelos anteriores, que los incrementos (disminuciones) elevados en la tasa desempleo provocan cambios más persistentes que incrementos (disminuciones) pequeños. En la figura 2.6 se presenta el logaritmo de la tasa de desempleo para EEUU.

Figura 2.8:

US Unemployment Rate, 1968:01-2003:12



Para contrastar la hipótesis de asimetrías en la persistencia del desempleo, planteamos la siguiente ecuación:

$$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \mu - \theta \varepsilon_{t-1} - \alpha(r) \varepsilon_{t-1} 1(|x_{t-1}| \neq r) + \varepsilon_t,$$

con x_t el incremento de la tasa de desempleo en el periodo t . Como puede verse en la ecuación, si $\alpha(r) \neq 0$, la persistencia de la perturbación dependerá del tamaño del incremento de la tasa de desempleo, representado por $|x_t|$. Dado que no existe una definición a priori de lo que es un incremento grande, esto es, no se conoce el valor de r^0 , usamos el estadístico propuesto en la Sección 2.5.2:

$$\sup_r \frac{T\hat{\alpha}^2(r)}{\hat{V}(\hat{\alpha}^2(r))}.$$

A través de los resultados del Teorema 2.6 y la aproximación bootstrap propuesta en dicha sección, realizamos el contraste de existencia de efecto threshold. El estadístico obtenido es 47.476, con un p -valor de 0.02. Por lo tanto, al 5 % rechazamos la hipótesis nula de simetría en la persistencia de las perturbaciones y aceptamos la existencia de histéresis.

Con este resultado, ajustamos un modelo ARTIMA(2,1,1,1) a la serie. Para minimizar la función objetivo, esto es, la suma de los errores al cuadrado, usamos la subrutina DBCPOL de FORTRAN. Los resultados se encuentran en la Tabla 2.6. Nótese, que para que la estimación sea consistente, x_t ha de ser ϕ -mixing. Sin embargo, este resultado es difícil de demostrar, con lo que entrará como un supuesto en el resto de esta sección¹.

Como predicen la teorías sobre los miembros y la duración, los incrementos pequeños serán menos persistentes que los grandes. Sin embargo, dado que la variable threshold es $|x_t|$, la histéresis es reversible. Los incrementos pequeños serán aquellos que supongan una variación inferior a 0.73 % de la tasa de desempleo. Esto supone el 28.18 % de los

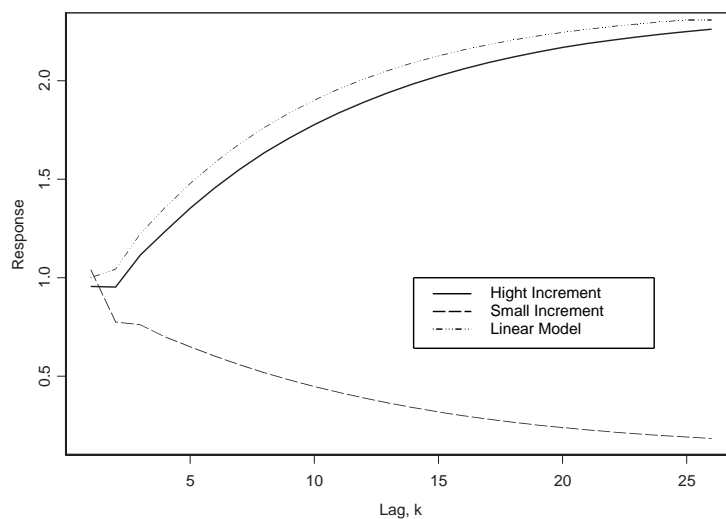
¹En cualquier caso, podría probarse que una condición suficiente para la consistencia es que x_t sea L_2 -NED de tamaño $-\infty$ en ε_t , (ver Anexo 1).

Cuadro 2.6: Resultados de la Estimación de los Modelos ARIMA y ARTIMA-Observable para la Tasa de Desempleo

	ARIMA(2,1,1)	ARTIMA(2,1,1,1) No restringido	ARTIMA(2,1,1,1) Restringido
μ	0.001 (0.003)	0.0001 (0.0003)	0.0001 (0.0003)
ρ_1	0.712 (0.097)	0.628 (0.047)	0.635 (0.048)
ρ_2	0.149 (0.058)	0.169 (0.046)	0.170 (0.046)
θ	0.669 (0.090)		
θ_1		0.582 (0.044)	0.588 (0.045)
θ_2		1.075 (0.070)	1
r		0.0073	0.0073
$\sigma_\varepsilon \times 10^2$	2.650	2.613	2.613
AIK	-4.413	-4.425	-4.429

incrementos. Además, no puede rechazarse al 5 % la existencia de una raíz unitaria en el régimen de los incrementos pequeños. Sin embargo, en principio no podemos inferir que las perturbaciones producidas en este régimen sean transitorias, ya que el modelo no cumple el supuesto I.1, que garantizaría que dichas perturbaciones son transitorias. Por cumplirse I.0, la variación provocada por la perturbación en dicho régimen se cancelará en el largo plazo. Sin embargo, esta perturbación afectará a la probabilidad de que en el periodo siguiente

Figura 2.9: Función Respuesta a un Impulso en la Tasa de Desempleo



se de un incremento grande, lo que afectará en media y de forma permanente a la serie. Por eso, la GI para la perturbación en este régimen no se anula. Pero este efecto no se descarta en el modelo de Blanchard y Summers o Ball, por lo que, aunque se hable en ellos de perturbaciones transitorias, este término no es del todo correcto bajo el prisma de la GI.

En la Figura 2.6 puede verse la función GI estimada para el modelo ARTIMA ajustado. La función GI estimada se ha dividido entre la perturbación para eliminar las distorsiones producidas por el tamaño. Claramente, las perturbaciones producidas en el régimen de incrementos pequeños son menos persistentes, y muy cercanas a las perturbaciones transitorias, ya que la GI desciende hacia cero. Esto implica que las perturbaciones afectan muy poco a la probabilidad de que se de cada régimen en el periodo siguiente.

2.7. Conclusiones

En este Capítulo se ha presentado un nuevo tipo de modelos no lineales con threshold en el componente de media móvil, y que denominamos ARTIMA. Estos modelos permiten generalizar las asimetrías en la persistencia que pueden presentar las perturbaciones de la serie. Como se prueba a lo largo del Capítulo, bajo determinadas condiciones descritas en el mismo, las perturbaciones podrán ser transitorias o permanentes dependiendo de alguna propiedad determinada por la variable threshold. Dependiendo de las características de esta variable threshold, distinguiremos dos tipos de modelos, ARTIMA-Observable (cuando la variable threshold es observable) y ARTIMA-Shock (cuando es la propia perturbación).

Igualmente, se profundiza en el estudio del modelo ARTIMA-Observable, obteniéndose las condiciones suficientes para su invertibilidad. A su vez, se propone como método de estimación el de mínimos cuadrados condicionales, probándose la consistencia e identificabilidad de los parámetros. También se obtiene la normalidad asintótica de todos los estimadores, excepto el del parámetro threshold. Al igual que en Chan (1993), se obtiene que el parámetro threshold, r , es superconsistente, y su distribución independiente de la del resto de estimadores.

De las propiedades básicas del modelo, dos son las hipótesis de interés: la existencia de asimetrías en la persistencia; y la posibilidad de que esta implique la existencia de perturbaciones transitorias y permanentes. Por lo tanto, se propone contrastar en una primera etapa la existencia de asimetrías. Aquí se distinguen dos casos, según r sea conocido o no. En el primer caso, utilizaremos un estadístico de Wald, cuya distribución asintótica será una χ^2 . En el segundo caso, r no estará identificado, utilizándose un estadístico basado en el supremo, y cuya distribución asintótica, aunque conocida, no será tabulable. Por ello, para obtener el p-valor se propone un método Wild Bootstrap. Para confirmar estos resultados, se realiza un pequeño estudio de Monte Carlo en el que se estudia el comportamiento de los estadísticos propuestos en muestras finitas.

En una segunda etapa, y cuando se rechace la nula de linealidad, podremos contrastar la existencia de perturbaciones transitorias y permanentes. Esta hipótesis se traduce en la existencia de una raíz unitaria en algún régimen de la media móvil. De nuevo, se propone un estadístico de Wald, cuya distribución asintótica es una χ^2 .

Finalmente, y a modo de ejemplo que ilustre el uso de los métodos propuestos, se estudia la hipótesis de la histéresis, propuesta en los modelos de Blanchard y Summers (1986) y Ball (1999), en la tasa de desempleo de EEUU a través de un modelo ARTIMA-Observable y con el incremento de la tasa como variable threshold. Se obtiene una respuesta afirmativa, esto es, la tasa de desempleo de equilibrio depende de su pasado, y además, depende de forma asimétrica.

2.8. Bibliografía

Andrews, D.W.K (1988): “Laws of large numbers for dependent non-identically distributed random variables”. *Econometric Theory*, Vol. **4**, 458-467.

Andrews, D.W.K. (1993): “Tests for Parameter Instability and Structural Change With Unknown Change Point”. *Econometrica*, Vol. **61**, 821-856.

Andrews, D.W.K. y Ploberger, W. (1994): “Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative”. *Econometrica*, Vol. **62**, 1383-1414.

Ball, L. (1999): “Aggregate demand and long-run unemployment”. *Brookings Papers on Economic Activity*, N 2, 189-251.

Beaudry, P. y Koop, G. (1994): “Do recessions permanently change output?”. *Journal of Monetary Economics*, Vol. **31**, 149-163.

Bec, F., Carrasco, M. y Salem M.B. (2004): “Tests for unit-root versus Threshold specification with an application to the PPP”. *Journal of Business & Economic Statistics*, Vol. **22**, 382-395.

Bierens, H. (1983): “Uniform consistency of kernel estimators of a regression function under generalized conditions”. *Journal of the American Statistical Association*, Vol. **77**, 699-707.

Billingsley, P. (1968): *Convergence of Probability Measures*. John Wiley.

Blanchard, O.J. y Summers, L.H. (1985): “Hysteresis and the european unemployment problem”. *NBER Working Paper*, N 1950.

Brännas, K. y Gooijer, J.G.D. (1994): “Autoregressive-asymmetric moving average models for business cycle data”. *Journal of Forecasting*, Vol. **13**, 529-544.

Caner, M. y Hansen, B.E. (2001): “Threshold Autoregression with a Unit Root”. *Econometrica*, Vol. **69**, 1555-1596.

Chan, K.S. (1993): "Consistency and limiting distribution of the least squares estimator of a Threshold Autoregressive Model". *The Annals of Statistics*, Vol. **21**, 520-533.

Davies, R.B. (1977): "Hypothesis testing when a nuisance parameter is present only under the alternative". *Biometrika*, Vol. **64**, 247-54.

Davies, R.B. (1987): "Hypothesis testing when a nuisance parameter is present only under the alternative". *Biometrika*, Vol. **74**, 33-43.

Davidson, J. (1994): *Stochastic Limit Theory*. Oxford University Press.

Doukhan, P. (1995): *Mixing, Properties and Examples*. Springer-Verlag.

Elwood, S.K. (1998): "Is the persistence of shocks to output asymmetric?". *Journal of Monetary Economics*, Vol. **41**, 411-426.

Engle, R.F. y Smith, A.D. (1999): "Stochastic Permanent Breaks". *The Review of Economics and Statistics*, Vol. **81**, 553-574.

Falk, B. (1986): "Further evidence on the asymmetric behavior of economic time series over the Business Cycle". *Journal of Political Economy*, Vol. **94**, 1096-1109.

Gallant, A.R. y White, H. (1988): *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Basil Blackwell.

González, M. y Gonzálo, J. (1998): "Threshold Unit Root Processes". Working Paper.

Gooijer, J.G. (1992): "On Threshold Moving-average Models". *Journal of Time Series Analysis*, Vol. **19**, 1-18.

Granger, C.W.J y Andersen, A. (1978): "On the invertibility of time series models". *Stochastic Processes and their Applications*, Vol. **8**, 87-92.

Granger, C.W.J. y Teräsvirta, T. (1993): *Modeling Nonlinear Economic Relationships*, Oxford University Press.

Guay, A. y Scaillet, O. (2003): “Indirect Inference, Nuisance Parameter, and Threshold Moving Average Models”. *Journal of Business and Economic Statistics*, Vol. **21**, 122-32.

Hallin, M. (1980): “Invertibility and generalized invertibility of time series models”. *J. R. Statist. Soc. B*, Vol. **42**, 210-212.

Hamilton, J. (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle”. *Econometrica* Vol. **75**, N2, 357-384.

Hansen, B.E. (1996): “Inference when a nuisance parameter is not identified under the null hypothesis”. *Econometrica*, Vol. **64**, 413-430.

Hansen, B.E. (2000): “Sample splitting and threshold estimation”. *Econometrica*, Vol. **68**, 575-603.

Ibragimov, I.A. (1962): “Some limit theorems for stationary processes”. *Theory of Probability and its Applications*, Vol. **7**, 349-3482.

Iosifescu, M. y Theodorescu, R. (1969): *Random Processes and Learning* . Springer-Verlag.

Koop, G., Pesaran, M.H. y Potter, S.M. (1996): “Impulse response analysis in nonlinear multivariate models”. *Journal of Econometrics*, Vol. **74**, 119-147.

Luukkonen, R., Saikkonen, P. y Teräsvirta, T. (1988): “Testing linearity against Smooth Transition Autoregressive Models”. *Biometrika*, Vol. **75**, 491 499.

McLeish, D. L. (1975): “A maximal inequality and dependent strong laws”. *Annals of Probability*, Vol. **3**, 826-836.

McLeish, D. L. (1977): “On the invariance principle for nonstationay mixingales”. *Annals of Probability*, Vol. **5**, 616-621.

Mokkadem, A. (1990): “Propriétés de mélange des modèles autorégressifs polynomiaux”. *Ann. I.H.P.*, Vol. **26**, 219-260.

Neftçi, S.N. (1984): “Are economic time series asymmetric over the business cycle?”. *Journal of Political Economy*, Vol. **92**, N 2, 307-328.

Obstfeld, M. y Taylor, A.M. (1997): “Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher’s commodity points revisited”. *Journal of the Japanese and International Economies*, Vol. **11**, 441-479.

Potter, S.M. (2000): “Nonlinear impulse response functions”. *Journal of Economic Dynamics & Control*, Vol. **24**, 1425-1446.

Pötscher, B.M. y Prucha, I.R. (1991): “Basic structure of the asymptotic theory in dynamic nonlinear econometric models, Part I: Consistency and approximation concepts”. *Econometric Reviews*, Vol. **10**, 125-216.

Pötscher, B.M. y Prucha, I.R. (1997): *Dynamic Nonlinear Econometric Models: Asymptotic Theory*, Springer-Verlag Berlin and Heidelberg GmbH & Co. K.

Rosenblatt, M. (1956): “A central limit theorem and a strong mixing condition”. *Proc. Nat. Acad. Sci. U.S.A*, Vol. **42**, 43-47.

Teräsvirta, T. (1994): “Specification, estimation, and evaluation of Smooth Transition Autoregressive Models”. *J. Amer. Statist. Assoc.*, Vol. **89**, 208218.

Tong, H. (1983): *Threshold Models in Non-linear Time Series Analysis*, Lecture Notes in Statistics, N 21. Springer, Heidelberg.

Wecker, W.E. (1981): “Asymmetric time series”. *Journal of the American Statistical Association*, Vol. **76**, N 373, 16-21.

White, H. (1984): *Asymptotic Theory for Econometricians*. Academic Press, Inc.

2.9. Anexo 1

Con la intención de facilitar la lectura de la Tesis y dotarla de un formato autocontenido, en esta sección se presentan los principales conceptos de dependencia de los que se hará uso en las demostraciones a lo largo de la Tesis. Junto con las definiciones, también se hará referencia a los principales resultados necesarios para la obtención de leyes de los grandes números. Estos se presentan en una serie de Teoremas propuestos por diferentes autores, y cuyas pruebas no reproduciremos aquí.

Los conceptos de dependencia de un proceso, x_t , definido en un espacio probabilístico $(\Omega, \mathfrak{F}, P)$, y que describiremos a continuación, se basan en los conjuntos de Borel generados por subconjuntos de la historia del proceso, y que denotaremos como $\mathfrak{F}_n^{n+m} = \sigma(x_n, \dots, x_{n+m})$. A continuación, definimos dos medidas de dependencia entre dos σ -álgebras.

Definition 2.7 Sean G y H dos σ -álgebras, entonces

$$\begin{aligned}\phi(G, H) &\equiv \sup_{\{g \in G, h \in H: P(g) > 0\}} |P(H|G) - P(H)| \\ \alpha(G, H) &\equiv \sup_{\{g \in G, h \in H\}} |P(H \cap G) - P(H)P(G)|.\end{aligned}$$

Basados en estas medidas, podemos definir los procesos ϕ -mixing o uniform mixing, (ver Iosifescu y Theodorescu, 1969) y α -mixing o strong mixing (ver Rosenblatt, 1956).

Definition 2.8 Sea x_t un proceso aleatorio, $\mathfrak{F}_{-\infty}^t$ la colección más pequeña de subconjuntos de Ω que contiene a la unión de σ -álgebras \mathfrak{F}_n^t cuando $n \rightarrow -\infty$, y $\mathfrak{F}_{t+m}^{\infty}$ la colección más pequeña de subconjuntos de Ω que contiene a la unión de σ -álgebras \mathfrak{F}_{t+m}^n cuando $n \rightarrow \infty$, definimos el coeficiente ϕ -mixing, $\phi(m) = \sup_t \phi(\mathfrak{F}_{-\infty}^t, \mathfrak{F}_{t+m}^{\infty})$ y α -mixing, $\alpha(m) = \sup_t \alpha(\mathfrak{F}_{-\infty}^t, \mathfrak{F}_{t+m}^{\infty})$. Por lo tanto, si $\phi(m) \rightarrow 0$ con $m \rightarrow \infty$, diremos que x_t es ϕ -mixing. Análogamente, si $\alpha(m) \rightarrow 0$ con $m \rightarrow \infty$, diremos que x_t es α -mixing.

Además, si $\phi(m) = O(m^{-a})$, x_t será ϕ -mixing de tamaño $-a$. Si $\alpha(m) = O(m^{-a})$, entonces x_t será α -mixing de tamaño $-a$.

Hay varias definiciones para el término “tamaño”, en esta Tesis adoptaremos la propuesta por Davidson (1994), McLeish (1975) y White (1984) entre otros, y que es la descrita en la definición 2.8. Entre la amplia literatura dedicada a los procesos mixing destacamos Doukhan (1995), que presenta un amplio número de propiedades y ejemplos, así como una detallada bibliografía al respecto. Como es fácilmente comprobable, ϕ -mixing es un concepto mixing más fuerte que α -mixing, en el sentido de que ϕ -mixing implica α -mixing, pero no al revés.

El problema fundamental de los procesos mixing aparece a nivel práctico, ya que puede implicar restricciones muy fuertes, (véase, por ejemplo, Davidson 1994, ejemplos 14.6 y 14.8). Si u_t es un proceso mixing, y x_t una función de infinitos retardos de u_t , $x_t = g(u_t, u_{t-1}, \dots)$, x_t puede no ser mixing. Por ello, en esta Tesis trabajaremos con un concepto de dependencia más potente a nivel práctico, introducido por McLeish (1975, 1977), que denominaremos mixingala y que supone una generalización del concepto de martingala en diferencias. Antes de definir el concepto de mixingala, definimos la norma L_p , que viene dada por

$$\|x_t\|_p = E^{1/p}(x_t^p).$$

Definition 2.9 Sea x_t una variable aleatoria integrable, diremos que es una L_p -mixingala si existe una secuencia de constantes no negativas $\{c_t\}_{-\infty}^{\infty}$ y $\{v_m\}_0^{\infty}$ tal que $v_m \rightarrow 0$ cuando $m \rightarrow \infty$ y

$$\|E(x_t | \mathfrak{F}_{t-m})\|_p \leq c_t v_m \tag{2.6}$$

$$\|x_t - E(x_t | \mathfrak{F}_{t+m})\|_p \leq c_t v_{m+1}, \tag{2.7}$$

se cumple para todo t y $m \geq 0$.

Claramente, una martingala en diferencias es una mixingala con $v_m = 0$ para $\forall m$. Además, si x_t es \mathfrak{S}_t medible, la condición 2.7 se cumple trivialmente. Análogamente a los procesos mixing, el tamaño de la mixingala vendrá determinado por el orden de convergencia de v_m , esto es, si $v_m = O(m^{-a})$, x_t será una L_2 - *mixingala* de orden $-a$ y constante c_t . Como en el caso de las martingalas, las desigualdades “maximales” son centrales en las aplicaciones del concepto de mixingala en la teoría asintótica y las leyes de los grandes números. El siguiente Teorema establece el principal resultado al respecto. Se trata de un corolario al Teorema 1.6 de McLeish (1975) y cuya prueba se puede encontrar en Davidson (1994) pág. 255.

Teorema 2.10 *Sea $\{x_t, \mathfrak{S}_t\}$ una secuencia L_2 - *mixingala* de tamaño $-1/2$ y constante c_t . Entonces*

$$E \left(\max_{1 \leq j \leq T} \left(\sum_{t=1}^T x_t \right)^2 \right) \leq K \sum_{t=1}^T c_t^2, \quad (2.8)$$

con $K < \infty$.

Lo que hace útil a nivel práctico el concepto de mixingala es la idea que se desarrolla en la siguiente definición, y que establece que aunque $x_t = g(u_t, u_{t-1}, \dots)$ no sea mixing, es suficiente con que dependa casi en su totalidad de los retardos más recientes de u_t . Esta idea fue introducida por Ibragimov (1962), y formalizada en diferentes definiciones por Billingsley (1968), McLeish (1975), Bierens (1983), Gallant y White (1988), Andrews (1988) y Pötscher y Prucha (1991), entre otros. La definición que se presenta en este anexo abarca y extiende la mayoría de las existentes, y se debe a Davidson (1994).

Definition 2.11 *Sea $\mathfrak{S}_{t-m}^{t+m} = \sigma(u_{t-m}, \dots, u_t, \dots, u_{t+m})$, para $\{u_t\}_{-\infty}^{\infty}$ una secuencia estocástica, posiblemente vectorial, en un espacio probabilístico $(\Omega, \mathfrak{S}, P)$. Si para $p > 0$ la secuencia de v.a integrables $\{x_t\}_{-\infty}^{\infty}$ satisface*

$$\|x_t - E(x_t | \mathfrak{S}_{t-m})\|_p \leq c_t v_m, \quad (2.9)$$

donde $v_m \rightarrow 0$ cuando $m \rightarrow \infty$ y $\{c_t\}_{-\infty}^{\infty}$ es una secuencia de constantes positivas, x_t será una secuencia near-epoch dependent en norma L_p ($L_p - NED$) en u_t .

Como en los dos casos anteriores, el tamaño vendrá dado por el orden de convergencia de v_m y su constante por c_t . En particular, si $v_m \sim m^\varphi \lambda^m$ con $\lambda < 1$ y $\varphi < \infty$, diremos que x_t es $L_p - NED$ en u_t de tamaño $-\infty$ y constante c_t . La ventaja de los procesos $L_p - NED$ es que desde el punto de vista práctico, engloban a una gran cantidad de modelos econométricos dinámicos. El supuesto de $L_p - NED$ no es exactamente un sustituto del de mixing, si no una propiedad del mapa de u_t a x_t . El concepto adquiere importancia, cuando u_t es mixing. Esto se pone de manifiesto en el siguiente teorema, que relaciona los conceptos de $L_p - NED$ y mixingala.

Teorema 2.12 Sea $\{x_t\}_{-\infty}^{\infty}$ una secuencia de media cero y L_r acotada, $\|x_t\|_r < \infty$, para $r > 1$. Sea $\{u_t\}_{-\infty}^{\infty}$ una secuencia ϕ -mixing, de tamaño $-a$. Si x_t es $L_p - NED$ de tamaño $-b$ en u_t para $1 \leq p \leq r$ con constantes $\{c_t\}_{-\infty}^{\infty}$, entonces $\{x_t, \mathfrak{F}_{-\infty}^t\}$ es una secuencia L_p -mixingala de tamaño $-\min\{b, a(1 - 1/r)\}$, con constantes $\bar{c}_t \ll \max\{c_t, \|x_t\|_r\}$.

El Teorema se puede relajar al caso en que u_t es α -mixing, pero entonces se ha de cumplir que $r > p$. Sin embargo, el caso $r = p$, con $p = 2$, resulta fundamental en las pruebas de las tasas de convergencia de los estimadores que se llevan a cabo en este capítulo y el siguiente. La prueba de este Teorema puede encontrarse en Davidson (1994) pág. 264.

Finalmente, presentamos dos Teoremas que establecen dos resultados básicos a la hora de trabajar con transformaciones de procesos $L_2 - NED$.

Teorema 2.13 Sean x_t e y_t procesos $L_p - NED$ en u_t de tamaños $-a_x$ y $-a_y$ y constantes $c_{x,t}$ y $c_{y,t}$, respectivamente. Entonces $x_t + y_t$ es $L_p - NED$ de tamaño $-\min\{a_x, a_y\}$ y constante $\max\{c_{x,t}, c_{y,t}\}$.

Teorema 2.14 Sean x_t e y_t procesos L_p -NED en u_t de tamaños $-a_x$ y $-a_y$ y constantes $c_{x,t}$ y $c_{y,t}$, respectivamente, tal que $\|x_t\|_{2r} < \infty$ y $\|y_t\|_{2r} < \infty$. Entonces $x_t y_t$ es L_2 -NED de tamaño $-\min\{a_x, a_y\}(r-2)/2(r-1)$, y constante v_{xy} , con

$$v_{xy}^{2(r-1)} = 2 \min \left\{ \left(c_{y,t}^{2-r} \|x_t\|_{2r}^2 \|y_t - E(y_t | \mathfrak{F}_{t-m})\|_{2r}^r \right), \left(c_{x,t}^{2-r} \|x_t\|_{2r}^2 \|x_t - E(x_t | \mathfrak{F}_{t-m})\|_{2r}^r \right) \right\}.$$

La prueba de estos dos Teoremas se puede encontrar en Davidson (1994), pág. 267 y pág. 273, respectivamente.

2.10. Anexo 2

Prueba del lema 2.1. Dado que $\Phi_p(L)$ tiene todas sus raíces fuera del círculo unidad, existirá $\Psi(L)$ con $\psi_0 = 1$ y $\psi_\infty = 0$. En primer lugar escribimos (2.2) de la siguiente manera,

$$\begin{aligned}
 (1-L)y_{t+k} &= \Psi(L)\mu + \Psi(L)(\varepsilon_{t+k} + \theta_{t+k-1}\varepsilon_{t+k-1}) \\
 &= \frac{\Psi(L)}{1-L}\mu + \sum_{l=0}^{\infty} \varepsilon_{t+k-l} + \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} (\psi_{j+1} - \theta_{t+k-1-j-l}\psi_j) \varepsilon_{t+k-1-j-l} \\
 &= \frac{\Psi(L)}{1-L}\mu + \varepsilon_{t+k} + \sum_{s=0}^{k-2} \left\{ \sum_{j=0}^s \psi_j (1 - \theta_{t+k-1-s}) \varepsilon_{t+k-1-s} + \psi_{s+1} \varepsilon_{t+k-1-s} \right\} + \\
 &\quad + \left[(1 - \theta_t) \sum_{j=0}^{k-1} \psi_j + \psi_k \right] \varepsilon_t + \\
 &\quad + \sum_{s=k}^{\infty} \left\{ \sum_{j=0}^s \psi_j (1 - \theta_{t+k-1-s}) \varepsilon_{t+k-1-s} + \psi_{s+1} \varepsilon_{t+k-1-s} \right\}.
 \end{aligned}$$

Usando I.1 y la ley de las esperanzas iteradas (LIE),

$$E(y_{t+k}|v_t, w_{t-1}) - E(y_{t+k}|w_{t-1}) = \left[(1 - \theta_t) \sum_{j=0}^{k-1} \psi_j + \psi_k \right] \varepsilon_t.$$

Finalmente, usando I.0,

$$\lim_{k \rightarrow \infty} GI(k, v_t, w_{t-1}) = \begin{cases} \psi_k \varepsilon_t & \text{if } \theta_t = 1 \\ \left[(1 - \theta) \sum_{j=0}^{k-1} \psi_j + \psi_k \right] \varepsilon_t & \text{if } \theta_t \neq 1. \end{cases}$$

■

Prueba del lema 2.2. Dado que z_t es observable, de la ecuación (2.2), $\hat{\varepsilon}_t$ y v_t (ver Definición 2.1) se pueden escribir como:

$$\begin{aligned}
 \hat{\varepsilon}_t &= \Phi_p(L)x_t - \mu + \theta_{t-1}(z_{t-1})\hat{\varepsilon}_{t-1}, \\
 v_t &= \hat{\varepsilon}_t - \varepsilon_t = \theta_{t-1}(z_{t-1})(\hat{\varepsilon}_{t-1} - \varepsilon_{t-1}), \\
 \text{con } \theta_{t-1}(z_{t-1}) &= \begin{cases} \theta_1 & \text{si } z_{t-1} > r \\ \theta_2 & \text{si } z_{t-1} \leq r. \end{cases}
 \end{aligned}$$

Por lo tanto es fácil ver que

$$v_t^2 = \prod_{j=1}^t \theta_{t-j}^2(z_{t-j}) v_0^2.$$

Una condición suficiente para la invertibilidad de y_t es,

$$\lim_{t \rightarrow \infty} E \left(\prod_{j=1}^t \theta_{t-j}^2(z_{t-j}) \right) = 0,$$

asumiendo que $v_0 < \infty$. Esta igualdad será cierta si los supuestos G.2 y A.0 se cumplen. Bajo estos supuestos, dado que $\theta_t(\cdot)$ es una función medible, $\theta_t(z_t)$ será ergódico. Usando el teorema de ergodicidad, $\frac{1}{T} \sum_{t=1}^T \ln \theta_t^2(z_t) \xrightarrow{as} E(\ln \theta_t^2(z_t)) < 0$. Por lo tanto, $\prod_{t=1}^T \theta_t^2(z_t) = \exp \left\{ \sum_{t=1}^T \ln \theta_t^2(z_t) \right\} \xrightarrow{as} 0$. Dada la integrabilidad uniforme de $\prod_{t=1}^T \theta_t^2(z_t)$, obtenemos que $\lim_{T \rightarrow \infty} \left\| \prod_{t=1}^T \theta_t^2(z_t) \right\|_1 = 0$, (ver el Teorema 18.14 de Davidson (1994)). ■

2.10.1. Consistencia y tasa de convergencia.

El Teorema 2.2 establece la tasa de convergencia del estimador CLS a los verdaderos valores de los parámetros del modelo ARTIMA-Observable, y por lo tanto, su identificabilidad y consistencia. Antes de probar la tasa de convergencia de dicho estimador, empezaremos por probar su identificabilidad y consistencia. Como se estableció en la Sección 2.4, la prueba se realiza para el caso $p = 1$, el caso más general es fácilmente derivable a partir de este caso particular. Por lo tanto, denotamos $\phi = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5) = (\theta_1, \theta_2, \rho, \mu, r)$. Análogamente, $\phi^0 = (\phi_1^0, \phi_2^0, \phi_3^0, \phi_4^0, \phi_5^0) = (\theta_1^0, \theta_2^0, \rho^0, \mu^0, r^0)$, denotarán el verdadero valor de los parámetros del proceso. Por el supuesto A.1 sabemos que $\phi \in \Theta$, si $(\theta_1, \theta_2, r) \in \Theta_\theta$, $\rho \in (-1, 1)$ y $\mu \in [-\bar{\mu}, \bar{\mu}]$. El estimador CLS de ϕ^0 , $\hat{\phi}_T$, vendrá dado por:

$$\hat{\phi}_T = \arg \min_{\phi \in \Theta} Q_T(\phi).$$

Siguiendo Pötcher and Prucha (1997) para probar la consistencia del estimador definiremos la siguiente distancia:

$$\varrho(\phi^*, \phi) = \max_i \{|\phi_i^* - \phi_i|\}.$$

Así, diremos que $\hat{\phi}_T$ es consistente si $\varrho(\hat{\phi}_T, \phi^0) \xrightarrow{p} 0$. Una condición suficiente para $\varrho(\hat{\phi}_T, \phi^0) \xrightarrow{p} 0$, es que para $\forall \epsilon > 0$

$$\liminf_{T \rightarrow \infty} P \left(\left[\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} Q_T(\phi) - Q_T(\phi^0) \right] > 0 \right) \xrightarrow{T \rightarrow \infty} 1. \quad (2.10)$$

Esta condición además, garantiza la identificabilidad de los parámetros ϕ^0 .

Por simplicidad, asumiremos que $\varepsilon_0 = \hat{\varepsilon}_0 = 0$, y por lo tanto $Q_T(\phi^0) = \sum_{t=1}^T \varepsilon_t^2$. Dada la invertibilidad del proceso bajo G.2 y A.1, los resultados se mantendrán si no se cumple dicha igualdad.

Así, es fácil probar que

$$\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} Q_T(\phi) - Q_T(\phi^0) \geq \inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} \sum_{t=1}^T v_t^2(\phi, \phi^0) - \sup_{\phi | \varrho(\phi, \phi^0) > \epsilon} \left| \sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0) \right|. \quad (2.11)$$

donde $v_t(\phi, \phi^0) = e_t(\phi) - \varepsilon_t$.

De la desigualdad (2.11) y la condición (2.10), es fácil ver, teniendo en cuenta que $T > 0$, que la consistencia de los estimadores se obtiene de las siguientes dos proposiciones.

Proposición 2.15 *Para todo $\epsilon > 0$ y $\eta > 0$,*

$$P \left[\sup_{\phi | \varrho(\phi, \phi^0) > \epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0)}{T} \right| < \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Proposición 2.16 *Para todo $\epsilon > 0$, $\exists a(\epsilon) > 0$ tal que*

$$P \left[\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} \frac{1}{T} \sum_{t=1}^T v_t^2(\phi, \phi^0) > a(\epsilon) \right] \xrightarrow{T \rightarrow \infty} 1.$$

Ambas proposiciones implican una ley uniforme de los grandes números (ULLN) para ambos procesos, $\varepsilon_t v_t(\phi, \phi^0)$ y $v_t^2(\phi, \phi^0)$. Para ello, será necesario acotar la variación de estos

procesos en ϕ con procesos que cumplan alguna ley de los grandes números. A continuación, definimos estos procesos “cota” que usaremos en las pruebas de ambas proposiciones:

$$\bar{e}_t = (|x_t| + |x_{t-1}| + \bar{\mu}) + \sum_{j=0}^{t-2} \prod_{i=0}^j |\bar{\theta}_{t-i-1}| (|x_{t-j-1}| + |x_{t-j-2}| + \bar{\mu}), \quad (2.12)$$

$$h_t^i(\phi') = \left. \frac{\partial e_t}{\partial \phi_i} \right|_{\phi'} \quad \text{para } i = 1, \dots, 4 \quad (2.13)$$

$$\bar{h}_t = [2\bar{e}_{t-1} + |x_{t-1}| + 1] + \sum_{j=0}^{t-2} \prod_{i=0}^j \bar{\theta}_{t-i-1} [2\bar{e}_{t-j-2} + |x_{t-1}| + 1] \quad (2.14)$$

$$\bar{\delta}_t(r^*, \nu) = 1(r^* - \nu < z_t < r^* + \nu) |2 + 2\delta|$$

$$v_t(\phi^*, \phi) = e_t(\phi^*) - e_t(\phi) \quad (2.15)$$

$$\bar{v}_t(\phi^*, \nu) = \nu [|x_{t-1}| + 1 + 2\bar{e}_{t-1}] + \delta_{t-1}(r^*, \nu) \bar{e}_{t-1} + \bar{\theta}_{t-1} \bar{v}_{t-1}(\phi^*, \phi) \quad (2.16)$$

$$s_t^{(i)}(\phi^*, \phi) = h_t^{(i)}(\phi^*) - h_t^{(i)}(\phi) \quad \text{para } i = 1, \dots, 4 \quad (2.17)$$

$$\bar{s}_t(\phi^*, \nu) = 2\delta_{t-1}(r^*, \nu) [\bar{e}_{t-1} + \bar{h}_{t-1}] + \bar{v}_{t-1}(\phi^*, \nu) + \nu \bar{h}_{t-1} + \bar{\theta}_{t-1} \bar{s}_{t-1}(\phi^*, \phi), \quad (2.18)$$

con $\nu > 0$. Como es fácil probar, $\sup_{\theta \in \Theta_\theta} |\theta_t(z_t)| \leq \bar{\theta}_t$, con $\bar{\theta}_t$ definido en el supuesto A.1. Con este resultado se tiene que $\sup_{\phi \in \Theta} |e_t(\phi)| \leq \bar{e}_t$ y el $\sup_{\phi \in \Theta} |h_t^{(i)}(\phi)| \leq \bar{h}_t$, donde \bar{e}_t y \bar{h}_t , no dependen de ϕ . Análogamente, el $\sup_{\phi/\varrho(\phi^*, \phi) \leq \nu} |v_t(\phi^*, \phi)| \leq \bar{v}_t(\phi^*, \nu)$ y $\sup_{\phi/\varrho(\phi^*, \phi) \leq \nu} |s_t(\phi^*, \phi)| \leq \bar{s}_t(\phi^*, \nu)$, donde $\bar{v}_t(\phi^*, \nu)$ y $\bar{s}_t(\phi^*, \nu)$ solo dependen de ϕ^* y ν , la distancia entre los parámetros. Así, estos procesos servirán de cota a las variaciones de sus respectivos procesos en ϕ . Además, como comentamos en un principio, es necesaria alguna ley de los grandes números (LLN) para cada uno de los procesos de (2.12) a (2.18). Claramente, los procesos de interés son complejos y no lineales, por lo que sus propiedades de memoria se establecerán en términos de $L_2 - NED$ y $L_2 - mixingalas$, introducidos por McLeish (1975), y cuya aplicación a este tipo de procesos es mucho más flexible. Una breve descripción y varios resultados de interés pueden encontrarse en el Anexo 1.

Los siguientes tres lemas establecen los momentos de los procesos (2.12) a (2.18), sus propiedades de memoria y una cota para la suma parcial de los mismos. Estos lemas serán

clave en la prueba de las proposiciones 2.15 y 2.16.

Lema 2.3 *Bajo los supuestos G.0 y A.1-A.2, para todo $0 < k \leq 2\gamma$, se cumple que*

$$(a) \quad E(\bar{e}_t^k) \leq K_e$$

$$(b) \quad E(\bar{h}_t^k) \leq K_h$$

$$(c) \quad E(\bar{v}_t^k(\phi^*, \nu)) \leq \max(\nu^k, \nu) K_v$$

$$(d) \quad E(\bar{s}_t^k(\phi^*, \nu)) \leq \max(\nu^k, \nu) K_s$$

con todas las constantes menores que ∞ .

Prueba del Lema 2.3. Usando la desigualdad de Minkowski, la ley de esperanzas iteradas y los supuestos del lema, se puede probar que para $t > \bar{q}$,

$$\begin{aligned} \|\bar{e}_t\|_k &\leq \|(|x_t| + |x_{t-1}| + \bar{\mu})\|_k + \sum_{j=1}^{t-2} \left\| \prod_{i=1}^j |\bar{\theta}_{t-i}| (|x_{t-j}| + |x_{t-j-1}| + \bar{\mu}) \right\|_k \\ &\leq \|6\varepsilon_t + \bar{\mu}\|_k \left[\sum_{n=0}^{\bar{q}-1} (1+\delta)^n \right] + \sum_{j=1}^{\lceil \frac{t-2}{\bar{q}} \rceil} \bar{\lambda}^j \|6\varepsilon_t + \bar{\mu}\|_k \left[\sum_{n=1}^{\bar{q}} (1+\delta)^n \right] + \\ &\quad + \bar{\lambda}^{\lceil \frac{t-2}{\bar{q}} \rceil} \sum_{n=1}^{t-2-\bar{q}\lceil \frac{t-2}{\bar{q}} \rceil} \|6\varepsilon_t + \bar{\mu}\|_k (1+\delta)^{nk} \\ &\leq \frac{K_{\delta, \bar{q}} \|6\varepsilon_t + \bar{\mu}\|_k}{1 - \bar{\lambda}} < \infty. \end{aligned}$$

De la misma manera, para \bar{h}_t tenemos que

$$\|\bar{h}_t\|_k \leq \frac{K_{\delta, \bar{q}} [\|\bar{e}_t\|_k + \|x_t\|_k + 1]}{1 - \bar{\lambda}} < \infty,$$

dado que $[\|\bar{e}_t\|_k + \|x_t\|_k + 1] < \infty$ por el resultado previo. Con lo que queda probado (a) y (b).

Usando la LIE, G.0 y A.2 es fácil probar que

$$E \left[|\bar{e}_{t-1}|^k \bar{\delta}_{t-1}^k(r^*, \nu) \right] \leq \nu K_e$$

donde K_e está acotada. Usando este resultado, y la desigualdad de Minkowsky, obtenemos que

$$\left\| \nu [|x_{t-1}| + 1 + 2\bar{e}_{t-1}] + \bar{\delta}_{t-1}(r^*, \nu) \bar{e}_{t-1} \right\|_k \leq \max(\nu, \nu^{1/k}) K_v^{1/k},$$

y siguiendo los mismos pasos que para $\|\bar{e}_t\|_k$, se prueba que

$$\|\bar{v}_t(\phi^*, \nu)\|_k \leq \max(\nu, \nu^{1/k}) \bar{K}_v^{1/k},$$

que implica (c). Para la prueba de (d) usamos además que

$$E \left[|\bar{h}_{t-1}|^k \bar{\delta}_{t-1}^k(r^*, \nu) \right] \leq \nu K_h,$$

el resto es similar a la de (c). ■

Lema 2.4 *Bajo los supuestos G.0, G.2 y A.1-A.2 \bar{e}_t , \bar{h}_t , $\bar{v}_t(\phi^*, \nu)$ y $\bar{s}_t(\phi^*, \nu)$ son procesos $L_{2\gamma}$ -acotados y L_2 -NED de tamaño $-\infty$ en $w_t = (\varepsilon_t, x_t, z_t)$ con constantes $\|\bar{e}_t\|_2$, $\|\bar{h}_t\|_2$, $\|\bar{v}_t(\phi^*, \nu)\|_2$ y $\|\bar{s}_t(\phi^*, \nu)\|_2$ respectivamente.*

Prueba del Lema 2.4. La primera parte de la prueba, $L_{2\gamma}$ -acotados, se obtiene directamente del lema 2.3.

Para probar que \bar{e}_t es L_2 -NED, (ver definición 2.11), debemos probar que

$$\|\bar{e}_t - E(\bar{e}_t / \mathfrak{S}_{t-m}^{t+m})\|_2 \leq d_t v_m, \quad (2.19)$$

donde $v_m = O(m^{-b})$, y d_t es una secuencia de constantes positivas finitas. Para probar (2.19), definimos

$$\bar{e}_t^m = (|x_t| + |x_{t-1}| + \bar{\mu}) + \sum_{j=1}^{m-1} \prod_{i=1}^j |\bar{\theta}_{t-i}| (|x_{t-j}| + |x_{t-j-1}| + \bar{\mu}),$$

que es \mathfrak{S}_{t-m}^{t+m} -medible. Al igual que en el lema 2.3, usando la LIE y el supuesto A.1,

$$\begin{aligned} \|\bar{e}_t - \bar{e}_t^m\|_2 &\leq \left\| \prod_{i=1}^{m-1} |\bar{\theta}_{t-i-1}| \sum_{j=m-1}^{t-2} \prod_{i=m-1}^j |\bar{\theta}_{t-i-1}| (|x_{t-j-1}| + |x_{t-j-2}| + \bar{\mu}) \right\|_2 \\ &\leq \bar{\lambda}^{\lceil \frac{m-1}{q} \rceil} \|\bar{e}_t\|_2. \end{aligned}$$

Dado que \bar{e}_t^m es una aproximación \mathfrak{S}_{t-m}^{t+m} - medible,

$$\|\bar{e}_t - E(\bar{e}_t / \mathfrak{S}_{t-m}^{t+m})\|_2 \leq \|\bar{e}_t - \bar{e}_t^m\|_2 \leq \bar{\lambda}^{\lceil \frac{m-2}{\bar{q}} \rceil} \|\bar{e}_t\|_2.$$

Finalmente, tomando $v_m^e = \bar{\lambda}^{\lceil \frac{m-2}{\bar{q}} \rceil}$ y $d_t^e = \|\bar{e}_t\|_2$, (2.19) queda probado, ya que v_m^e decrece exponencialmente con m y su tasa de convergencia será ∞ por definición. Para \bar{h}_t , definimos $m = m^*\bar{q} + 2$ con $m^* \geq 2$ una secuencia de enteros positivos, y

$$\bar{h}_t^{m^*\bar{q}+2} = \left[2\bar{e}_{t-1}^{m^*\bar{q}+1} + 1 \right] + \sum_{j=1}^{m^*\bar{q}} \prod_{i=1}^j \bar{\theta}_{t-i} \left[2\bar{e}_{t-j-1}^{m^*\bar{q}-j+1} + 1 \right].$$

Como en el caso de \bar{e}_t ,

$$\begin{aligned} \|\bar{h}_t - \bar{h}_t^{m^*\bar{q}+2}\|_2 &\leq 2 \left\| \bar{e}_{t-1} - \bar{e}_{t-1}^{m^*\bar{q}+1} \right\|_2 + 2 \sum_{j=1}^{m^*\bar{q}} \left\| \prod_{i=1}^j \bar{\theta}_{t-i} \left| \bar{e}_{t-j-1} - \bar{e}_{t-j-1}^{m^*\bar{q}-j+1} \right| \right\|_2 \\ &\quad + \left\| \prod_{i=1}^{m^*\bar{q}} |\bar{\theta}_{t-i-1}| \sum_{j=m^*\bar{q}+1}^{t-2} \prod_{i=m^*\bar{q}+1}^j |\bar{\theta}_{t-i}| (2\bar{e}_{t-j-1}) \right\|_2 \\ &\leq 2v_{m^*\bar{q}+1}^e d_t^e + 2 \sum_{j=1}^{m^*} \left[\sum_{i=1}^{\bar{q}-1} (1+\delta)^i v_{m^*\bar{q}-i-\bar{q}(j-1)+1}^e d_t^e \right] \bar{\lambda}^{j-1} + \bar{\lambda}^{m^*} \|\bar{h}_t\|_2 \\ &\leq \left[\bar{\lambda}^{m^*} + \sum_{j=1}^{m^*} \bar{\lambda}^{j-1} \bar{\lambda}^{m^*-j+1} + \bar{\lambda}^{m^*} \right] 2 \|\bar{h}_t\|_2 = (2 + m^*) \bar{\lambda}^{m^*} 2 \|\bar{h}_t\|_2, \end{aligned}$$

usando que $v_{m^*\bar{q}-i-\bar{q}(j-1)+1}^e = \bar{\lambda}^{\lceil \frac{m^*\bar{q}-i-\bar{q}(j-1)+1}{\bar{q}} \rceil} = \bar{\lambda}^{m^*-j+1} \bar{\lambda}^{-\lceil \frac{i-1}{\bar{q}} \rceil} = \bar{\lambda}^{m^*-j+1}$ dado que $i < \bar{q}$. Definiendo $v_m^h = \left(3 + \lceil \frac{m-2}{\bar{q}} \rceil \right) \bar{\lambda}^{\lceil \frac{m-2}{\bar{q}} \rceil}$ y $d_t^h = \|\bar{h}_t\|_2$, queda probado que \bar{h}_t es $L_2 - NED$ de tamaño $-\infty$. Para los otros dos casos, $\bar{v}_t(\phi^*, \nu)$ y $\bar{s}_t(\phi^*, \nu)$, la prueba sigue los mismos pasos que los dos casos anteriores. ■

Lema 2.5 *Bajo los supuestos G.0, G.2 y A.1-A.2,*

$$E \left[\max_{1 \leq j \leq T} \left(\sum_{t=1}^j \chi_t \right)^2 \right] \leq K \sum_{t=1}^T \|\chi_t\|_2^2,$$

para $\chi_t = \bar{e}_t - E(\bar{e}_t)$, $\bar{h}_t - E(\bar{h}_t)$, $|\varepsilon_t| \bar{v}_t - E(|\varepsilon_t| \bar{v}_t)$ y $|\varepsilon_t| \bar{s}_t - E(|\varepsilon_t| \bar{s}_t)$ con $K < \infty$.

Prueba del Lema 2.5. Dado que $|\varepsilon_t|$ es ϕ -mixing, es directo probar que $|\varepsilon_t| \bar{v}_t$ y $|\varepsilon_t| \bar{s}_t$ son $L_2 - NED$ con las mismas propiedades que \bar{v}_t y \bar{s}_t . Además, los cuatro procesos son $L_{2\gamma}$ -bounded, por lo que la prueba de este lema se obtiene directamente de la aplicación del lema 2.4 y los Teoremas 2.10 y 2.12 del Anexo 1. ■

Vistos los procesos y sus propiedades a través de estos tres lemas, pasamos a continuación a probar las proposiciones 2.15 y 2.16, que a su vez prueban la identificabilidad de los parámetros y consistencia de los estimadores.

Prueba de la Proposición 2.15. La proposición 2.15 establece que para todo $\epsilon > 0$ y $\eta > 0$

$$P \left[\sup_{\phi | \varrho(\phi, \phi^0) > \epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0)}{T} \right| < \eta \right] \xrightarrow{T \rightarrow \infty} 1,$$

donde $v_t(\phi, \phi^0)$ se define en (2.15). Para probar esta convergencia, descompondremos el proceso $v_t(\phi, \phi^0)$ en su parte diferenciable y no diferenciable. Para ello, usaremos que

$$v_t(\phi, \phi^0) = v_t(\phi, \phi') + v_t(\phi', \phi^0),$$

donde $\phi' = (\theta_1^0, \theta_2^0, \rho^0, \mu^0, r)$. Para un r dado, $v_t(\phi, \phi')$ es diferenciable con respecto a ϕ_i , para $i = 1, \dots, 4$ y usando un desarrollo de Taylor de orden uno en ϕ' , tenemos que

$$\sum_{t=1}^T v_t(\phi, \phi') = \sum_{t=1}^T \sum_{i=1}^4 (\phi_i - \phi_i^0) h_t^{(i)}(\phi^0) + \sum_{t=1}^T \sum_{i=1}^4 (\phi_i - \phi_i^0) \left(h_t^{(i)}(\phi^0) - h_t^{(i)}(\phi^*) \right), \quad (2.20)$$

donde $h_t^{(i)}$ se define en (2.13) y $\varrho(\phi', \phi^*) \leq \varrho(\phi', \phi)$.

Definimos $B_\epsilon = \{\phi \in \Theta | \varrho(\phi, \phi^0) > \epsilon\}$. Entonces,

$$\sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0)}{T} \right| \leq \sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi')}{T} \right| \quad (2.21)$$

$$+ \sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi', \phi^0)}{T} \right| \quad (2.22)$$

La parte diferenciable, $(2.21)^2$, usando la ecuación (2.20), se puede escribir como

$$\sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=2}^T \varepsilon_t v_t(\phi, \phi')}{T} \right| \leq \sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 (\phi_i - \phi_i^0) h_t^i(\phi^0)}{T} \right| + \quad (2.23)$$

$$+ \sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 (\phi_i - \phi_i^0) (h_t^i(\phi^0) - h_t^i(\phi))}{T} \right|. \quad (2.24)$$

Para probar la convergencia a 0 del término diferenciable, trabajamos con cada uno de los supremos de (2.23) y (2.24). Empezamos con (2.23). Dado que $\varepsilon_t h_t^{(i)}(\phi^0)$ no depende de ϕ , tenemos que $\sup_{\phi \in B_\epsilon} \left| T^{-1} \sum_{t=1}^T \varepsilon_t h_t^{(i)}(\phi^0) \right| = \left| T^{-1} \sum_{t=1}^T \varepsilon_t h_t^{(i)}(\phi^0) \right|$, que es una suma de diferencia de martingalas (mds), ya que $h_t^{(i)}(\phi^0)$ es \mathfrak{F}_{t-1} -medible y ε_t iid. Además $h_t^{(i)}(\phi^0) \leq \bar{h}_t$, con \bar{h}_t definida en (2.14) y L_2 -acotada, con lo que obtenemos que para $\eta > 0$,

$$P \left[\sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=1}^T \varepsilon_t h_t^{(i)}(\phi^0)}{T} \right| > \eta \right] \leq \frac{\sigma_\varepsilon^2 \left(\|h_t^{(i)}(\phi^0)\|_2 \right)^2}{T \eta^2} \xrightarrow{T \rightarrow \infty} 0.$$

Finalmente,

$$\sup_{\phi \in B_\epsilon} \left| \frac{2 \sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 (\phi_i - \phi_i^0) h_t^{(i)}(\phi^0)}{T} \right| \leq 2 |2\bar{\mu} + 6| \left| \frac{\sum_{t=1}^T \varepsilon_t h_t^{(i)}(\phi^0)}{T} \right|,$$

para $i = 1, 2, \dots, 4$, por lo tanto, queda probado que el término de (2.23) convergerá a 0.

El otro término, (2.24), sí depende ϕ . Para probar la convergencia acotaremos el proceso a través de una suma de procesos convergentes. Para ello usaremos que el espacio paramétrico para ϕ_l es acotado, de forma que podemos definir j_l y J_l para $l = 1, 2, \dots, 5$, con $j_l = -J_l, -J_l + 1, \dots, J_l$, tal que $\phi_l^{j_l} = \phi_l^0 + j_l \nu$ establece una partición de cada subespacio paramétrico, con $\nu > 0$ y que definiremos posteriormente. Basado en esta partición para cada uno de los parámetros, definiremos la siguiente secuencia de vectores de índices, $\ell_j = (j_1, j_2, j_3, j_4, j_5)$, para $j = 1, \dots, J$, y vectores de parámetros $\phi^{\ell_j} = (\phi_1^{j_1}, \phi_2^{j_2}, \phi_3^{j_3}, \phi_4^{j_4}, \phi_5^{j_5})$ con $\phi_l^{j_l}$ para $l = 1, 2, \dots, 5$ definido anteriormente. Finalmente, definiremos $B_{\epsilon, j_l}(\phi) = \left\{ \phi \in B_\epsilon \mid 0 \leq |\phi_l - \phi_l^{\ell_j}| \leq \nu \right\}$ y $B_{\epsilon, \ell_j}(\phi) = \prod_{l=1}^5 B_{\epsilon, j_l}(\phi)$ con $\prod_{l=1}^5$ el producto

²En general, y salvo que se exprese lo contrario, nos referiremos siempre a la parte derecha de la igualdad o desigualdad.

cartesiano de los 5 subconjuntos. Con estas definiciones, es fácil probar que $B_{\epsilon, \ell_j}(\phi)$ constituye una partición de B_ϵ , con $B_\epsilon \subset \cup_{\ell_j=1}^J B_{\epsilon, \ell_j}(\phi)$ y $J = \prod_{l=1}^5 J_l$. Como en el caso de (2.23),

$$\sup_{\phi \in B_\epsilon} \left| \frac{2 \sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 (\phi_i - \phi_i^0) (h_t^{(i)}(\phi^0) - h_t^{(i)}(\phi^*))}{T} \right| \leq \sup_{\phi \in B_\epsilon} 2 |\phi_i - \phi_i^0| \sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^*)}{T} \right|,$$

para $i = 1, 2, \dots, 4$, con s_t^i definido en (2.17) y $\sup_{\phi \in B_\epsilon} 2 |\phi_i - \phi_i^0| < 2|2\bar{\mu} + 6| < \infty$. Usando la partición $B_{\epsilon, \ell_j}(\phi)$ definida anteriormente, se tiene que

$$\sup_{\phi^* \in B_\epsilon} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^*)}{T} \right| \leq \max_{\ell_j} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T} \right| + \quad (2.25)$$

$$+ \max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi, \phi^{\ell_j})}{T} \right|. \quad (2.26)$$

Por lo tanto, para (2.25) y todo $\eta > 0$

$$P \left[\max_{\ell_j} \left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T} \right| > \eta \right] \leq \sum_{j=1}^J P \left[\left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T} \right| > \eta \right]. \quad (2.27)$$

Dado que $s_t^{(i)}(\phi^0, \phi^{\ell_j})$ es \mathfrak{F}_{t-1} -medible, $\varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j})$ es una mds con $E \left(\varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j}) \right)^2 = E(\varepsilon_t)^2 E \left(s_t^{(i)}(\phi^0, \phi^{\ell_j}) \right)^2$. Del lema 2.3, la esperanza de $s_t^i(\phi, \phi^{\ell_j})$ solo dependerá de $\varrho(\phi^0, \phi^{\ell_j})$, y por lo tanto del $\max_l |j_l|$. Con este resultado y usando la desigualdad de Markov, en la suma de (2.27) podemos distinguir cinco casos, dependiendo de que índice en $|j_l|$ sea el máximo, obteniendo que para todo $\eta > 0$,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T} \right| > \eta \right] &\leq 5 \sum_{j=1}^{J^*} \frac{j^5 \max(\nu^2, \nu) \sigma_\varepsilon^2 K_s}{T \eta^2} \leq \\ &\leq \frac{K_1}{T \eta^2 \nu^5} \xrightarrow{T \rightarrow \infty} 0, \end{aligned}$$

con $J^* = \max_l J_l$. Para el otro término, (2.26), usamos que $\sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} \left| s_t^{(i)}(\phi, \phi^{\ell_j}) \right| \leq$

$\bar{s}_t(\phi^{\ell_j}, \nu)$, donde \bar{s}_t está definido en (2.18), para obtener que

$$\sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi, \phi^{\ell_j})}{T} \right| \leq \frac{4 \sum_{t=1}^T (|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu) - E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu)])}{T} + \frac{\sum_{t=1}^T E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu)]}{T}.$$

Así, usando los lemas 2.3 (d) y 2.5 se puede probar que para todo $\eta > 0$,

$$P \left[\max_{\ell_j} \frac{\sum_{t=1}^T |\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu) - E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu)]}{T} > \eta/2 \right] \leq \frac{K_2}{T\eta^2},$$

y

$$\max_{\ell_j} \frac{\sum_{t=1}^T E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu)]}{T} \leq \nu K_s$$

Tomando ν , que define la partición B_{ϵ, ℓ_j} , suficientemente pequeño, tal que $\nu K_s < \frac{\eta}{2}$, tenemos que para todo $\eta > 0$,

$$P \left[\max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi, \phi^{\ell_j})}{T} \right| > \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Dado que $\sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} |\phi_i - \phi_i^0| < 2|2\bar{\mu} + 6| < \infty$ y los resultados de (2.25) y (2.26), para todo $\eta > 0$,

$$P \left[\sup_{\phi \in B_{\epsilon}} \left| \frac{\sum_{t=2}^T \varepsilon_t \sum_{i=1}^4 (\phi_i - \phi_i^0) s_t^{(i)}(\phi^0, \phi)}{T} \right| > \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Con este resultado, y los anteriores obtenidos para (2.23) se prueba que el término diferenciable, (2.21), converge a 0.

En el término no diferenciable, (2.22), sólo $\phi_5 = r$ puede ser diferente, luego usaremos la siguiente partición, $B_{\epsilon, j}(\phi) = \{\phi \in B_{\epsilon} / \phi_5 \in (r_0 + \nu(j+1), r_0 + \nu j)\}$,

con $\phi'^j = (\phi_1^0, \phi_2^0, \phi_3^0, \phi_4^0, r_0 + \nu j)$ y $j = -J_5, \dots, J_5$, para acotar la suma de forma similar a (2.24),

$$\sup_{\phi \in B_{\epsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t(\phi', \phi^0)}{T} \right| \leq \max_j \left| \frac{\sum_{t=1}^T \varepsilon_t v_t(\phi'^j, \phi^0)}{T} \right| \quad (2.28)$$

$$+ \max_j \sup_{\phi \in B_{\epsilon, j}(\phi)} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t(\phi, \phi'^j)}{T} \right|. \quad (2.29)$$

Como en los casos anteriores, probamos que cada término converge en probabilidad a 0. Para el primer término, (2.28), el lema 2.3 (c) junto con que $v_t(\phi'^j, \phi^0)$ es \mathfrak{F}_{t-1} -medible, garantizan que $\varepsilon_t v_t(\phi'^j, \phi^0)$ es una *mds* con $E(\varepsilon_t^2 v_t^2(\phi'^j, \phi^0)) \leq \nu K_v$. Luego es fácil probar que

$$P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi'^j, \phi^0)}{T} \right| > \eta \right] \leq \frac{K_c}{\nu T \eta^2} \xrightarrow{T \rightarrow \infty} 0.$$

Para el segundo término, (2.29), la prueba será similar a la de (2.26). En este caso, $\sup_{B_{\epsilon,j}(\phi)} \varepsilon_t v_t(\phi, \phi'^j) \leq |\varepsilon_t| \bar{v}_t(\phi'^{j+1}, \phi'^j)$ con lo que por los lemas 2.3 (c) y 2.5 se prueba que para todo $\eta > 0$,

$$P \left[\max_j \sup_{B_{\epsilon,j}(\phi)} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t(\phi, \phi'^j)}{T} \right| > \frac{\eta}{2} \right] \leq \frac{K_d}{T \eta^2} \xrightarrow{T \rightarrow \infty} 0,$$

siguiendo los mismos pasos que en (2.26).

Este resultado termina la prueba de la proposición 2.15. La convergencia en probabilidad a 0 de los términos (2.28) y (2.29) implica la convergencia en probabilidad a 0 de (2.22). Igualmente, la convergencia en probabilidad a 0 de (2.23) y (2.24), implica lo mismo para (2.21). Con esto, queda probado que para $\forall \eta > 0$,

$$P \left[\sup_{\phi \in B_\epsilon} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0)}{T} \right| \geq \eta \right] \xrightarrow{T \rightarrow \infty} 0,$$

y con ello la proposición 2.15. ■

Prueba de la Proposición 2.16. La proposición 2.16 establece que, para todo $\epsilon > 0$, $\exists a(\epsilon) > 0$ tal que

$$P \left[\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} \frac{1}{T} \sum_{t=1}^T v_t^2(\phi, \phi^0) > a(\epsilon) \right] \xrightarrow{T \rightarrow \infty} 1,$$

o lo que es lo mismo,

$$P \left[\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) > a(\epsilon) \right] \xrightarrow{T \rightarrow \infty} 1 \quad (2.30)$$

con,

$$\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} \sum_{\iota=1}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) = \inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} \sum_{\iota=1}^{\lfloor \frac{T}{2} \rfloor} v_{k(\iota)-1}^2(\phi, \phi^0) + v_{k(\iota)}^2(\phi, \phi^0), \quad (2.31)$$

para $k(\iota) = 2\iota$ y $[x]$ la parte entera de x . Definiendo $\delta_{\phi, l}^0 = \phi_l - \phi_l^0$, podemos expresar $v_{k(\iota)}(\phi, \phi^0)$ como,

$$v_{k(\iota)}(\phi, \phi^0) = \delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1} + \theta_{k(\iota)-1} v_{k(\iota)-1}(\phi, \phi^0),$$

y condicionando al suceso

$$E = \left\{ |v_{k(\iota)-1}(\phi, \phi^0)| > \frac{1}{2(1+\delta)} |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1}| \right\},$$

y a su complementario, E^c , obtenemos que

$$W_{k(\iota)}(\phi, \phi^0) \geq \frac{1}{4(1+\delta)^2} |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1}|^2 = \frac{1}{4(1+\delta)^2} \vartheta_{k(\iota)}^2(\delta_\phi^0)$$

con

$$\begin{aligned} \vartheta_{k(\iota)}^2(\delta_\phi^0) &= [\vartheta_{k(\iota), 1}^2(\delta_\phi^0) + \vartheta_{k(\iota), 2}^2(\delta_\phi^0)] \\ \vartheta_{k(\iota), 1}^2(\delta_\phi^0) &= 1(z_{k(\iota)-1} < \underline{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 2}^0 \varepsilon_{k(\iota)-1}|^2 + \\ &\quad + 1(z_{k(\iota)-1} > \bar{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 1}^0 \varepsilon_{k(\iota)-1}|^2 \\ \vartheta_{k(\iota), 2}^2(\delta_\phi^0) &= 1(r^0 < z_{k(\iota)-1} < r^0 + \delta_{\phi, 5}^0) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_1^0 - \theta_2) \varepsilon_{k(\iota)-1}|^2, \end{aligned}$$

si $\delta_{\phi, 5}^0 > 0$. El caso $\delta_{\phi, 5}^0 < 0$ es idéntico. Además, distinguiremos otros dos casos, cuando $\max_{i \leq 4} |\delta_{\phi, i}^0| > b_1$ para un valor de b_1 , $0 < b_1 < \epsilon$, que definiremos más adelante, y su complementario. Definiendo

$$\begin{aligned} M_{k(\iota)-1}(\underline{r}) &= 1(z_{k(\iota)-1} < \underline{r}) (x_{k(\iota)-1}, 1, \varepsilon_{k(\iota)-1})' (x_{k(\iota)-1}, 1, \varepsilon_{k(\iota)-1}) \\ M_{k(\iota)-1}(\bar{r}) &= 1(z_{k(\iota)-1} > \bar{r}) (x_{k(\iota)-1}, 1, \varepsilon_{k(\iota)-1})' (x_{k(\iota)-1}, 1, \varepsilon_{k(\iota)-1}), \end{aligned}$$

y usando el supuesto A.3 se obtiene que,

$$E [\vartheta_{k(\iota),1}^2 (\delta_\phi^0)] = c_1^{0'} E (M_{k(\iota)-1}(\bar{r})) c_1^0 + c_2^{0'} E (M_{k(\iota)-1}(\underline{r})) c_2^0 \geq \max_{i \leq 4} |\delta_{\phi,i}^0|^2 a_1, \quad (2.32)$$

con $a_1 > 0$, $c_1^0 = (\delta_{\phi,4}^0, \delta_{\phi,3}^0, \delta_{\phi,1}^0)$ y $c_2^0 = (\delta_{\phi,4}^0, \delta_{\phi,3}^0, \delta_{\phi,2}^0)$. Usando además que $M_{k(\iota)-1}(\underline{r})$ y $M_{k(\iota)-1}(\bar{r})$ son ϕ -mixing y c_1^0 y c_2^0 acotados es fácil probar que,

$$\begin{aligned} P \left[\sup_{c_1^0} \frac{1}{T} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} c_1^{0'} [M_{k(\iota)-1}(\bar{r}) - E(M_{k(\iota)-1}(\bar{r}))] c_1^0 \right| > \eta \right] &\xrightarrow{T \rightarrow \infty} 0 \\ P \left[\sup_{c_2^0} \frac{1}{T} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} c_2^{0'} [M_{k(\iota)-1}(\underline{r}) - E(M_{k(\iota)-1}(\underline{r}))] c_2^0 \right| > \eta \right] &\xrightarrow{T \rightarrow \infty} 0. \end{aligned}$$

Por lo tanto, si $\max_{i \leq 4} |\delta_{\phi,i}^0| > b_1$, tenemos que

$$P \left[\inf_{\phi | \varrho(\phi, \phi^0) > \epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),1}^2 (\delta_\phi^0) > E [\vartheta_{k(\iota),1}^2 (\delta_\phi^0)] - \eta \right] \xrightarrow{T \rightarrow \infty} 1, \quad (2.33)$$

con $E [\vartheta_{k(\iota),1}^2 (\delta_\phi^0)] - \eta > 0$ por (2.32) y para η suficientemente pequeño.

Para el segundo caso, $\max_{i \leq 4} |\delta_{\phi,i}^0| < b_1$, probamos que

$$P \left[\sup_{r/r-r^0 > \epsilon} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2)] \right| > \eta \right] \leq \epsilon \quad (2.34)$$

$$P \left[\sup_{r/r-r^0 > \epsilon} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^0, r) N_{k(\iota)-1} - E(\delta_{k(\iota)}(r^0, r) N_{k(\iota)-1})] \right| > \eta \right] \leq \epsilon, \quad (2.35)$$

con $\delta_{k(\iota)}(r, r') = 1$ ($r < z_{k(\iota)-1} < r'$) y $N_{k(\iota)-1} = |x_{k(\iota)-1} \varepsilon_{k(\iota)-1} + \varepsilon_{k(\iota)-1}|$. Únicamente probaremos la desigualdad de (2.34), ya que la otra desigualdad es similar. Para ello, usamos

la partición $B_{\epsilon,j}(\phi)$, definida en la proposición 2.15, de forma que

$$\begin{aligned} \sup_{r|r-r^0>\epsilon} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2)] \right| \leq \\ \max_j T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^0, r^j) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^0, r^j) \varepsilon_{k(\iota)-1}^2)] \right| + \\ + \max_j T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2)] \right| + \\ + \max_j \sup_{r \in B_{\epsilon,j}(\phi)} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [E(\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2) - E(\delta_{k(\iota)}(r^j, r) \varepsilon_{k(\iota)-1}^2)] \right|, \end{aligned}$$

usando que $\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2$ es creciente en r . Usando los resultados del lema 2.3 y que $\varepsilon_{k(\iota)-1}^2$ es ϕ -mixing, siguiendo los mismos pasos que en la ecuación (2.28), es fácil obtener que para todo $\eta > 0$

$$\begin{aligned} P \left[\max_j T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^0, r^j) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^0, r^j) \varepsilon_{k(\iota)-1}^2)] \right| > \eta \right] &\leq \frac{K}{T\nu\eta^2} \xrightarrow{T \rightarrow \infty} 0 \\ P \left[\max_j T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2 - E(\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2)] \right| > \eta \right] &\leq \frac{K}{T\eta^2} \xrightarrow{T \rightarrow \infty} 0. \end{aligned}$$

Además,

$$\max_j \sup_{r \in B_{\epsilon,j}(\phi)} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} [E(\delta_{k(\iota)}(r^j, r^{j+1}) \varepsilon_{k(\iota)-1}^2) - E(\delta_{k(\iota)}(r^j, r) \varepsilon_{k(\iota)-1}^2)] \right| \leq \nu K,$$

con νK tan pequeño como queramos. Así, estos tres últimos resultados prueban la desigualdad (2.34). Siguiendo los mismos pasos, obtenemos la desigualdad (2.35). Con estas dos desigualdades, y asumiendo $b_1 > 0$ suficientemente pequeño, se prueba que

$$P \left[\inf_{r|r-r^0>\epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \delta_{k(\iota)}(r^0, r) \left[(\delta^0 - b_1)^2 \varepsilon_{k(\iota)-1}^2 - 2(\delta^0 - b_1) b_1 N_{k(\iota)-1} \right] > \epsilon a_2 \right] \xrightarrow{T \rightarrow \infty} 1,$$

con $a_2 > 0$. Por lo tanto, se tiene que,

$$\inf_{r/|r-r^0|>\epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\delta_\phi^0) \geq$$

$$\inf_{r/|r-r^0|>\epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \delta_{k(\iota)}(r^0, r) \left[(\delta^0 - b_1)^2 \varepsilon_{k(\iota)-1}^2 - 2(\delta^0 - b_1) b_1 N_{k(\iota)-1} \right],$$

y

$$P \left[\inf_{r/|r-r^0|>\epsilon} T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\delta_\phi^0) > \epsilon a_2 \right] \xrightarrow{T \rightarrow \infty} 1. \quad (2.36)$$

Los resultados (2.33) y (2.36) prueban la proposición 2.16. ■

Probadas ambas proposiciones, queda probada la consistencia de los estimadores CLS y la identificabilidad de los parámetros.

A continuación desarrollamos la prueba respecto a la tasa de convergencia. Esta se basa en la prueba del Teorema 2 de Chan (1993).

Prueba del Teorema 2.2.

Para el Teorema 2.2 debemos probar que $\widehat{\phi}_{l,T} = \phi_l^0 + O_p(T^{-1/2})$ para $l = 1, \dots, 4$ y $\widehat{r}_T = r_0 + O_p(T^{-1})$. Definimos, por tanto,

$$B_{\Delta,\epsilon} = \left\{ \phi \in \Theta : \sqrt{\frac{\Delta}{T}} \leq |\phi_l - \phi_l^0| = \delta_{\phi,l}^0 \leq \epsilon \text{ para } l = 1, \dots, 4 \text{ y } \frac{\Delta}{T} \leq |r - r^0| = \delta_{\phi,5}^0 \leq \epsilon \right\},$$

que será acotado. La prueba se desarrolla para algún $\epsilon > 0$ suficientemente pequeño. Por la consistencia de los estimadores, $\widehat{\phi}_T \in B_{\Delta,\epsilon}$ con probabilidad tan cercana a 1 como queramos, para todo $T \geq T_\epsilon^*$, con T_ϵ^* suficientemente grande. Probar que la tasa de convergencia es la establecida en el Teorema 2.2 será lo mismo que probar que para todo $\varepsilon > 0$ y $T \geq T_\varepsilon^*$ existe $\Delta < \infty$ tal que $P \left[\max_{i \leq 4} |\widehat{\phi}_{i,T} - \phi_i^0| > \sqrt{\frac{\Delta}{T}} \right] < \varepsilon$ y $P \left[|\widehat{r}_T - r^0| > \frac{\Delta}{T} \right] < \varepsilon$. Como en el caso de consistencia, una condición suficiente es que para todo $\varepsilon > 0$ y $T \geq T_\varepsilon^*$ existe $\Delta < \infty$ tal que:

$$P \left[\inf_{\phi \in B_{\Delta,\epsilon}} Q_T(\omega, \phi) - Q_T(\omega, \phi^0) > 0 \right] \geq 1 - \varepsilon,$$

o lo que es lo mismo,

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{Q_T(\omega, \theta) - Q_T(\omega, \theta^0)}{T\bar{\varrho}^2(\phi, \phi^0)} > 0 \right] \geq 1 - \varepsilon, \quad (2.37)$$

con $\bar{\varrho}(\phi, \phi^*) = \max_{l \leq 4} (|\phi_l - \phi_l^*|, |\phi_5 - \phi_5^*|^{1/2})$, dado que $T\bar{\varrho}^2(\phi, \phi^0) > 0$. Los pasos para probar (2.37), son los mismos que para la consistencia, aunque en este caso, las sumas se estandarizan con $[T\bar{\varrho}^2(\phi, \phi^0)]^{-1}$ en vez de T^{-1} . Así, (2.37) se deriva de las siguientes dos proposiciones, similares a las proposiciones 2.15 y 2.16.

Proposición 2.17 *Para todo $\eta, \varepsilon > 0$, $\exists \Delta < \infty$ tal que*

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t(v_t(\phi, \phi^0))}{T\bar{\varrho}^2(\phi, \phi^0)} \right| > \eta \right] \leq \varepsilon.$$

Proposición 2.18 *Para todo $\varepsilon > 0$, $\exists \Delta < \infty$ y $\eta > 0$ tal que*

$$P \left[\inf_{\phi \in B_{\Delta, \epsilon}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T\bar{\varrho}^2(\phi, \phi^0)} > \eta \right] \geq 1 - \varepsilon.$$

Prueba de la Proposición 2.17. A lo largo de esta prueba, η y ε serán dos números reales positivos cualesquiera. El proceso de interés, $\varepsilon_t v_t(\phi, \phi^0)$, es el mismo que el de la proposición 2.15, luego sus propiedades serán las mismas. Como en aquel caso, descomponemos el proceso en su parte diferenciable y no diferenciable. En la parte diferenciable, debemos probar que, para algún Δ suficientemente grande,

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi')}{T\bar{\varrho}^2(\phi, \phi^0)} \right| > \eta \right] \leq \varepsilon. \quad (2.38)$$

Usando un desarrollo de Taylor de orden uno en ϕ' , obtenemos los términos correspondientes a (2.23) y (2.24). Así, para el término correspondiente a (2.23), se tiene que

$$\begin{aligned} \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=2}^T \varepsilon_t (\phi_i - \phi_{(i)}^0) h_t^{(i)}(\phi^0)}{T\bar{\varrho}^2(\phi, \phi^0)} \right| &\leq \left| \frac{\sum_{t=2}^T \varepsilon_t h_t^{(i)}(\phi^0)}{\sqrt{T\Delta}} \right| \\ P \left[\left| \frac{\sum_{t=2}^T \varepsilon_t h_t^{(i)}(\phi^0)}{\sqrt{T\Delta}} \right| > \eta \right] &\leq \frac{\sigma_\varepsilon^2 \left(\|h_t^{(i)}(\phi^0)\|_2 \right)^2}{\Delta \eta^2}, \end{aligned} \quad (2.39)$$

usando que $\sup_{\phi \in B_{\Delta, \epsilon}} \frac{|\phi_i - \phi_i^0|}{\bar{\varrho}^2(\phi, \phi^0)} \leq \frac{1}{\bar{\varrho}(\phi, \phi^0)}$ y $\varepsilon_t h_t^{(i)}(\phi^0)$ es una *mds*. Tomando Δ suficientemente grande, la probabilidad de (2.39) será tan pequeña como queramos.

Para el otro término, (2.24), desarrollamos la prueba para el caso $\phi_l > \phi_l^0$, con $l = 1, \dots, 5$. El resto de casos se obtienen siguiendo los mismos pasos. De forma similar a la proposición 2.15, definimos $\phi_l^{j_l} = \phi_l^0 + \sqrt{\frac{\Delta}{T} b^{j_l}}$ para $l = 1, 2, \dots, 4$, $\phi_5^{j_5} = r^{j_5} = r^0 + \frac{\Delta}{T} b^{j_5}$ y $b > 1$. Para cada subespacio paramétrico, definimos la siguiente partición, $B_{\epsilon, j_l}^l(\Delta, \phi_l) = \left\{ \phi_l \in B_{\Delta, \epsilon}^+ \mid 0 \leq \frac{\phi_l - \phi_l^{j_l}}{\phi_l^{j_l} - \phi_l^0} \leq (b-1) \right\}$, con $B_{\Delta, \epsilon}^+ = B_{\Delta, \epsilon}(\phi_l^{j_l} > \phi_l^0)$. Como entonces, definiendo, $\ell_j = (j_1, j_2, j_3, j_4, j_5)$, para $j = 1, \dots, J$, es directo probar que $B_{\Delta, \epsilon}^+ \subset \cup_{\ell_j=1}^J B_{\epsilon, \ell_j}^{\Delta}(\phi)$, con $B_{\epsilon, \ell_j}^{\Delta}(\phi) = \prod_{l=1}^5 B_{\epsilon, j_l}^l(\Delta, \phi_l)$. Igualmente, se puede probar que para $\phi^{\ell_j} = (\phi_1^{j_1}, \phi_2^{j_2}, \phi_3^{j_3}, \phi_4^{j_4}, \phi_5^{j_5})$, el $\sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} \bar{\varrho}(\phi^{\ell_j}, \phi) = \nu_{\Delta}(\ell_j) = \bar{\varrho}(\phi^{\ell_j}, \phi^0) (b-1)^{1/2}$.

Por lo tanto, tenemos que,

$$\sup_{\phi \in B_{\Delta, \epsilon}^+} \left| \frac{\sum_{t=2}^T (\phi_i - \phi_i^0) \varepsilon_t s_t^{(i)}(\phi^0, \phi)}{T \bar{\varrho}^2(\phi, \phi^0)} \right| \leq \epsilon \max_{\ell_j} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T \bar{\varrho}^2(\phi^{\ell_j}, \phi^0)} \right| +$$

$$\max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} \left(\left| \frac{(\phi_i^{\ell_j} - \phi_i^0) \sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T \bar{\varrho}^2(\phi^{\ell_j}, \phi^0)} - \frac{(\phi_i - \phi_i^0) \sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi)}{T \bar{\varrho}^2(\phi, \phi^0)} \right| \right),$$

dado que $\max_{\ell_j} |\phi^{\ell_j} - \phi_i^0| < \epsilon$. Además, usando que

$$\max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} \left[\frac{\bar{\varrho}^2(\phi^{\ell_j}, \phi^0) \frac{(\phi_i - \phi_i^0)}{\phi_i^{\ell_j} - \phi_i^0} - \bar{\varrho}^2(\phi, \phi^0)}{\bar{\varrho}^2(\phi, \phi^0)} \right] \leq 2,$$

obtenemos que

$$\sup_{\phi \in B_{\Delta, \epsilon}^+} \left| \frac{\sum_{t=2}^T (\phi_i - \phi_i^0) \varepsilon_t s_t^{(i)}(\phi^0, \phi)}{T \bar{\varrho}^2(\phi, \phi^0)} \right| \leq \max_{\ell_j} 3 \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^{(i)}(\phi^0, \phi^{\ell_j})}{T \bar{\varrho}^2(\phi^{\ell_j}, \phi^0)} \right| +$$

$$+ \max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} 3 \left| \frac{\sum_{t=2}^T \varepsilon_t \bar{s}_t(\phi^{\ell_j}, \nu(\ell_j))}{T \bar{\varrho}(\phi^{\ell_j}, \phi^0)} \right|. \quad (2.40)$$

Para el primer sumando de (2.40), como en el caso del sumando (2.27), distinguimos cinco casos dependiendo de que componente de ℓ_j sea máximo. Usando la teoría sobre *mds*

obtenemos que,

$$P \left[\max_{\ell_j} \left| \frac{\sum_{t=2}^T \varepsilon_t s_t^i(\phi^0, \phi^{\ell_j})}{T \bar{\varrho}^2(\phi^{\ell_j}, \phi^0)} \right| > \eta \right] \leq 5 \sum_{j=1}^{J^*} \frac{j^4 \left(\frac{\Delta}{T} b^j\right) \sigma_\varepsilon^2 K_s^1}{T \eta^2 \left(\frac{\Delta}{T} b^j\right)^2} \leq \sum_{j=1}^{\infty} \frac{j^4 K'_a}{\eta^2 \Delta b^j}, \quad (2.41)$$

con $J^* = \max_l J_l$. Tomando $b > 1$, $\sum_{j=1}^{\infty} \frac{j^4 K'_a}{\eta^2 \Delta b^j} < \infty$, con lo que para Δ suficientemente grande la probabilidad será tan pequeña como queramos. Para el segundo sumando de (2.40), siguiendo los mismos pasos que en (2.41), obtenemos que

$$P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T |\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu(\ell_j)) - E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu(\ell_j))]}{T \bar{\varrho}(\phi^{\ell_j}, \phi^0)} \right| > \eta \right] \leq \sum_{j=1}^{J^*} \frac{j^4 K_b}{\eta^2 \Delta b^j} \quad (2.42)$$

$$\max_{\ell_j} \frac{E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu(\ell_j))]}{\bar{\varrho}(\phi^{\ell_j}, \phi^0)} \leq (b-1)^{1/2} K_c, \quad (2.43)$$

donde la última desigualdad, (2.43), se obtiene por el lema 2.3, que garantiza que

$E[|\varepsilon_t| \bar{s}_t(\phi^{\ell_j}, \nu(\ell_j))] \leq \nu(\ell_j) K_s \leq K_s \bar{\varrho}(\phi^{\ell_j}, \phi^0) (b-1)^{1/2}$. Tomando b suficientemente cercano a 1, esta cantidad será tan pequeña como queramos, con lo que para Δ suficientemente grande obtenemos el resultado deseado para el segundo término de (2.40). Con este resultado, concluimos la prueba para el caso en que el $\sup_{\phi \in B_{\Delta, \epsilon}}$ del término diferenciable pertenece a $B_{\Delta, \epsilon}^+$. El resto de casos, se obtiene de idéntica forma.

La parte no diferenciable, sólo depende del término r , nos concentramos en $r > r^0$, (el otro caso es igual), y por tanto, definimos la partición $B_{\Delta, j}(\phi) = \{\phi_i = \phi_i^0, r \in (r^j, r^{j+1}), i \leq 4\}$ con $r^j = r^0 + \frac{\Delta}{T} b^j$ y $\sup_j \frac{\Delta}{T} b^j \leq \epsilon$. Sea $\delta_{\phi, 5}^j = |r^j - r^0|$ y $\nu_j = (r^{j+1} - r^j) = \frac{\Delta}{T} b^j (b-1)$, debemos probar que,

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi', \phi^0)}{T \bar{\varrho}^2(\phi^{\ell_j}, \phi^0)} \right| > \eta \right] \leq P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi', \phi^0)}{T \delta_{\phi, 5}^j} \right| > \eta \right] \leq \varepsilon. \quad (2.44)$$

Como en los casos anteriores, obtenemos los siguientes resultados,

$$P \left[\max_j \left| \frac{\sum_{t=1}^T 2\varepsilon_t v_t(\phi'^j, \phi^0)}{T \delta_{\phi, 5}^j} \right| > \eta \right] \leq \frac{1}{\Delta} \sum_{j=1}^{\infty} \frac{K_d}{b^j \eta^2} \quad (2.45)$$

$$P \left[\max_j \left| \frac{\sum_{t=1}^T |\varepsilon_t| \bar{v}_t(\theta'^j, \nu_j) - E(|\varepsilon_t| \bar{v}_t(\theta'^j, \nu_j))}{T \delta_{\phi,5}^j} \right| > \eta \right] \leq \frac{1}{\Delta} \sum_{j=1}^{\infty} \frac{(b-1) K_d}{b^j \eta^2} \quad (2.46)$$

$$\max_j \frac{E(|\varepsilon_t| \bar{v}_t(\theta'^j, \nu_j))}{\delta_{\phi,5}^j} \leq (b-1) K'_d \quad (2.47)$$

$$\max_j \sup_{\phi \in B_{\Delta,j}(\phi)} \frac{\delta_{\phi,5}^j}{\delta_{\phi,5}} \leq b, \quad (2.48)$$

que conjuntamente, y para Δ suficientemente grande y b cercano a 1, prueban (2.44), y por lo tanto la proposición 2.17. ■

Prueba de la Proposición 2.18. La prueba de esta proposición es paralela a la proposición 2.16. En concreto, debemos probar que para algún Δ suficientemente grande y algún $\eta > 0$,

$$P \left[\inf_{\phi \in B_{\Delta,\epsilon}} \sum_{t=1}^T \frac{(v_t(\phi, \phi^0))^2}{T \bar{\varrho}^2(\phi, \phi^0)} > \eta \right] \geq 1 - \varepsilon,$$

para $\forall \varepsilon > 0$.

De la proposición 2.16,

$$\inf_{\phi \in B_{\Delta,\epsilon}} \sum_{t=1}^T \frac{(v_t(\phi, \phi^0))^2}{T \bar{\varrho}^2(\phi, \phi^0)} \geq \inf_{\phi \in B_{\Delta,\epsilon}} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\vartheta_{k(\iota),1}^2(\delta_\phi^0)}{T \bar{\varrho}^2(\phi, \phi^0)} + \inf_{\phi \in B_{\Delta,\epsilon}} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\vartheta_{k(\iota),2}^2(\delta_\phi^0)}{T \bar{\varrho}^2(\phi, \phi^0)}.$$

Como en aquel caso, dividimos $B_{\Delta,\epsilon}$ en dos subespacios. El primero, cuando el máximo de $|\phi_l - \phi_l^0|$ se obtiene para $l = 5$, lo cual implica que $\bar{\varrho}^2(\phi, \phi^0) = |r - r^0| > \frac{\Delta}{T}$. Bajo este supuesto, bastará probar que para todo $\varepsilon > 0$, existe Δ suficientemente grande y $\eta > 0$ tal que

$$P \left[\inf_{\phi \in B_{\Delta,\epsilon}} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\vartheta_{k(\iota),2}^2(\delta_\phi^0)}{T |r - r^0|} > \eta \right] \geq 1 - \varepsilon. \quad (2.49)$$

Como entonces, este resultado se deriva de los siguientes dos resultados,

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2 - E \left(\delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2 \right)}{T |r - r^0|} \right| > \eta \right] \leq \varepsilon \quad (2.50)$$

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\delta_{k(\iota)}(r^0, r) N_{k(\iota)-1} - E \left(\delta_{k(\iota)}(r^0, r) N_{k(\iota)-1} \right)}{T |r - r^0|} \right| > \eta \right] \leq \varepsilon, \quad (2.51)$$

ya que $\inf_{\phi \in B_{\Delta, \epsilon}} E \left(1 \left(r^0 < z_{k(\iota)-1} < r^0 + \delta_{\phi, 5}^0 \right) \varepsilon_{k(\iota)-1}^2 \right) / |r - r^0| > \underline{\sigma}_{\varepsilon/z}^2 > 0$ y $\sup_{\phi \in B_{\Delta, \epsilon}} E \left(\delta_{k(\iota)}(r^0, r) N_{k(\iota)-1} \right) / |r - r^0| \leq \bar{\sigma}_{\varepsilon/z}^2 M$, con lo que para ϵ y η suficientemente pequeños,

$$P \left[\inf_{\phi \in B_{\Delta, \epsilon}} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{(\delta^0 - \epsilon)^2 \delta_{k(\iota)}(r^0, r) \varepsilon_{k(\iota)-1}^2 - (\delta^0 - \epsilon) \epsilon \delta_{k(\iota)}(r^0, r) N_{k(\iota)-1}}{T |r - r^0|} > a_1 \right] \geq 1 - \varepsilon,$$

con $a_1 = (\delta^0 - \epsilon)^2 \underline{\sigma}_{\varepsilon/z}^2 - (\delta^0 - \epsilon) \epsilon \bar{\sigma}_{\varepsilon/z}^2 M - \eta > 0$. Para probar las desigualdades de (2.50) y (2.51), usaremos la partición basada en $r_j = r^0 \pm \frac{\Delta}{T} b^j$ con $b > 1$, como en la proposición 2.17, y seguiremos los pasos seguidos en (2.34).

Para el segundo caso, cuando el máximo de $|\phi_l - \phi_l^0|$ se alcanza para $l \leq 4$, tenemos que

$$\begin{aligned} \inf_{\phi \in B_{\Delta, \epsilon}} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\vartheta_{k(\iota), 1}^2(\delta_\phi^0)}{T \bar{\varrho}^2(\phi, \phi^0)} &\geq \inf_c \frac{1}{T} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} c' M_{k(\iota)-1}(\bar{r}) c \\ \inf_{\phi \in B_{\Delta, \epsilon}} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\vartheta_{k(\iota), 1}^2(\delta_\phi^0)}{T \bar{\varrho}^2(\phi, \phi^0)} &\geq \inf_c \frac{1}{T} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} c' M_{k(\iota)-1}(\underline{r}) c, \end{aligned}$$

con $c \in [1, 1]^3$ y tal que $\max_i c_i^2 = 1$. Como en la proposición 2.16, usando que $M_{k(\iota)-1}(\bar{r})$

y $M_{k(\iota)-1}(\underline{r})$ son ϕ -mixing y $\max_{i,j} c_i c_j \leq 1$, es fácil probar que

$$P \left[\sup_c \frac{1}{T} \left| \sum_{\iota=2}^{\left[\frac{T}{2}\right]} c' [M_{k(\iota)-1}(\bar{r}) - E(M_{k(\iota)-1}(\bar{r}))] c \right| > \eta \right] \xrightarrow{T \rightarrow \infty} 0$$

$$P \left[\sup_c \frac{1}{T} \left| \sum_{\iota=2}^{\left[\frac{T}{2}\right]} c' [M_{k(\iota)-1}(\underline{r}) - E(M_{k(\iota)-1}(\underline{r}))] c \right| > \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Usando que $E(M_{k(\iota)-1}(\bar{r}))$ y $E(M_{k(\iota)-1}(\underline{r}))$ son definidas positivas y $\max_i c_i^2 = 1$, obtenemos que

$$c' [E(M_{k(\iota)-1}(\bar{r}))] c \geq a_2 + \eta > 0$$

$$c' [E(M_{k(\iota)-1}(\underline{r}))] c \geq a_3 + \eta > 0.$$

Por lo tanto, para T y Δ suficientemente grandes,

$$P \left[\sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\vartheta_{k(\iota),2}^2(\delta_\phi^0)}{T\bar{\varrho}^2(\phi, \phi^0)} \geq a_1 \right] \geq 1 - \varepsilon \quad \text{si } \bar{\varrho}^2(\phi, \phi^0) = |r - r^0|$$

$$P \left[\inf_{\phi \in B_{\Delta, \varepsilon}} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\vartheta_{k(\iota),1}^2(\delta_\phi^0)}{T\bar{\varrho}^2(\phi, \phi^0)} \geq a_2 \right] \geq 1 - \varepsilon \quad \text{si } \bar{\varrho}^2(\phi, \phi^0) = |\phi_l - \phi_l^0| \text{ para } l = 1, 3, 4$$

$$P \left[\inf_{\phi \in B_{\Delta, \varepsilon}} \sum_{\iota=2}^{\left[\frac{T}{2}\right]} \frac{\vartheta_{k(\iota),1}^2(\delta_\phi^0)}{T\bar{\varrho}^2(\phi, \phi^0)} \geq a_3 \right] \geq 1 - \varepsilon \quad \text{si } \bar{\varrho}^2(\phi, \phi^0) = |\phi_l - \phi_l^0| \text{ para } l = 2, 3, 4.$$

Así,

$$P \left[\inf_{\phi \in B_{\Delta, \varepsilon}} \sum_{t=1}^T \frac{(v_t(\phi, \phi^0))^2}{T\bar{\varrho}^2(\phi, \phi^0)} > \min_{i \leq 3} (a_i) \right] \geq 1 - \varepsilon,$$

con $\min_{i \leq 3} (a_i) > 0$. Con esto, la proposición 2.18 queda probada. ■

2.10.2. Inferencia.

Normalidad asintótica cuando r^0 es conocido.

A lo largo de esta sección, dado que r^0 es conocido, trabajaremos con $\beta = (\beta_1, \beta_2, \beta_3, \beta_4) = (\theta_1, \theta_2, \rho, \mu)$ y $\beta^0 = (\theta_1^0, \theta_2^0, \rho^0, \mu^0)$. Además, todos los procesos que en su definición dependen de r , serán evaluados en r^0 , sin que se especifique explícitamente en la notación. Como vimos en la Sección 2.4, la función objetivo será diferenciable respecto a β . De la prueba del Teorema 2.2 sabemos que

$$Q_T(\beta) - Q_T(\beta^0) = \sum_{t=1}^T (v_t(\beta, \beta^0))^2 + \sum_{t=1}^T 2\varepsilon_t v_t(\beta, \beta^0).$$

Si definimos

$$2H_t^{i,l}(\beta^*) = \frac{\partial^2 (v_t(\beta, \beta^0))^2}{\partial \beta_i \partial \beta_l} \Big|_{\beta=\beta^*} = 2 \left[h_t^{(l)}(\beta^*) h_t^{(i)}(\beta^*) + v_t(\beta^*, \beta^0) h_t^{i,l}(\beta^*) \right],$$

$$h_t^{i,l}(\beta^*) = \frac{\partial^2 e_t(\beta)}{\partial \beta_i \partial \beta_l} \Big|_{\beta=\beta^*} = \frac{\partial h_t^{(i)}(\beta)}{\partial \beta_l} \Big|_{\beta=\beta^*},$$

$Q_T(\hat{\beta}) - Q_T(\beta^0)$ se podrá escribir, utilizando un desarrollo de Taylor de orden uno en torno a $Q_T(\beta^0)$, como

$$Q_T(\hat{\beta}) - Q_T(\beta^0) = \frac{1}{2} \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\hat{\beta}_i - \beta_i^0) 2H_t^{i,l}(\beta^*) (\hat{\beta}_l - \beta_l^0) + \sum_{i=1}^4 \sum_{t=1}^T 2(\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\beta^0) + \frac{1}{2} \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T 2(\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{il}(\beta^*) (\hat{\beta}_l - \beta_l^0), \quad (2.52)$$

con β^* tal que $\varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)$, $\hat{\beta} = \arg \min_{\beta \in \Theta_\beta} Q_T(\beta)$, $\Theta_\beta = \{\phi = (\beta, r^0) \in \Theta\}$ y usando que $\frac{\partial (v_t(\beta, \beta^0))^2}{\partial \beta_i} \Big|_{\beta^0} = 0$.

Para obtener la normalidad asintótica usaremos los resultados de los tres lemas que se describen a continuación, y que muestran el comportamiento asintótico de cada uno de los tres sumandos de la ecuación (2.52).

Lema 2.6 *Bajo los supuestos G.0, G.2, A.1, A.2 y A.3,*

$$\sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t h_t^{il}(\beta^*) \left(\hat{\beta}_l - \beta_l^0 \right) = o_p(1)$$

Prueba del Lema 2.6. Será suficiente probar que para todo $i, l \leq 4$ y $\eta > 0$

$$\lim_{T \rightarrow \infty} P \left[\sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_t h_t^{il}(\beta^*) \right| > \eta \right] = 0,$$

dado que, $\hat{\beta}$ es $O_p(T^{-1/2})$ y por lo tanto, $T \left(\hat{\beta}_i - \beta_i^0 \right) \left(\hat{\beta}_l - \beta_l^0 \right) = O_p(1)$. Este resultado se puede obtener fácilmente del Teorema 2.2 sin necesidad de asumir G.1, que sólo es necesario en la proposición 2.16 y 2.18 para la consistencia de \hat{r} . Dado que r^0 en este caso es conocido, G.1 no es necesario y la prueba se deriva directamente del Teorema 2.2.

De la definición de $h_t^{i,l}(\phi)$, y dado que $h_t^{i,l}(\phi)$ es derivable respecto a β , se puede probar que

$$\sup_{\beta / \varrho(\beta, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \left| h_t^{i,l}(\beta^0) - h_t^{i,l}(\beta) \right| \leq \varrho(\hat{\beta}, \beta^0) \bar{h}_t^{i,l}, \quad (2.53)$$

con $\bar{h}_t^{i,l} \mathfrak{S}_{t-1} - medible$ y $E \left[\left(\bar{h}_t^{i,l} \right)^2 \right] < \infty$, para $i, l \leq 4$. Por lo tanto,

$$\sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_t h_t^{il}(\beta^*) \right| \leq \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_t h_t^{il}(\beta^0) \right| + \varrho(\hat{\beta}, \beta^0) \frac{1}{T} \sum_{t=1}^T |\varepsilon_t| \bar{h}_t^{i,l}.$$

Dado que $\varepsilon_t h_t^{il}(\beta^0)$ es una mds , $E \left(\left| h_t^{i,l}(\beta^0) \right| \right) \leq K_{hh}$, $\varrho(\hat{\beta}, \beta^0) = o_p(1)$ y $\frac{1}{T} \sum_{t=1}^T |\varepsilon_t| \bar{h}_t^{i,l} = O_p(1)$, el lema queda probado.

■

Lema 2.7 *Bajo los supuestos G.0, G.2, A.1, A.2 y A.3,*

$$\begin{aligned} & \frac{1}{T} \sum_{t=1}^T \left(H_t^{i,l}(\beta^0) - H^{i,l}(\beta^0) \right) = o_p(1) \\ & \sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) \left[H_t^{i,l}(\beta^*) - H_t^{i,l}(\beta^0) \right] \left(\hat{\beta}_l - \beta_l^0 \right) = o_p(1), \end{aligned}$$

con $H^{i,l}(\beta^0) = E \left(h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0) \right)$.

Prueba del Lema 2.7. La primera parte del lema, dado que $H_t^{i,l}(\beta^0) = h_t^{(i)}(\beta^0)h_t^{(l)}(\beta^0)$ es $L_2 - mixingala$ de tamaño $-1/2$ para $i, j \leq 4$, se obtiene directamente de la aplicación de una LLN para $L_2 - mixingalas$, (ver por ejemplo McLeish, 1975, o el Corolario 20.16 de Davidson, 1994). Esta propiedad de $H_t^{i,l}(\beta^0)$ se obtiene siguiendo los mismos argumentos de los lemas 2.4 y 2.5.

Para la segunda parte, por definición

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \left(H_t^{i,l}(\beta) - H_t^{i,l}(\beta^0) \right) &= \frac{1}{T} \sum_{t=1}^T s_t^{(i)}(\beta, \beta^0) h_t^{(l)}(\beta^0) + \frac{1}{T} \sum_{t=1}^T s_t^{(l)}(\beta, \beta^0) h_t^{(i)}(\beta) \\ &\quad + \frac{1}{T} \sum_{t=1}^T v_t(\beta, \beta^0) h_t^{i,l}(\beta). \end{aligned}$$

Por lo tanto, el $\sup_{\beta^*/\varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T \left(H_t^{i,l}(\beta) - H_t^{i,l}(\beta^0) \right) \xrightarrow{p} 0$, si para cada uno de estos sumandos su supremo converge a 0. Primero probamos que para todo $\eta > 0$, se cumple que

$$\lim_{T \rightarrow \infty} P \left[\sup_{\beta^*/\varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T s_{t-1}^{(i)}(\beta, \beta^0) h_{t-1}^{(l)}(\beta) > \eta \right] = 0.$$

De las definiciones (2.14) y (2.18), y usando que $\phi_5 = \phi_5^0$, se obtiene que $\bar{s}_t(\beta^0, \varrho(\beta, \beta^0)) = \varrho(\beta, \beta^0) \bar{s}_t$, donde \bar{s}_t no depende de ϕ y $E(\bar{s}_t^4) \leq K_s < \infty$. Así

$$\sup_{\beta^*/\varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T s_t^{(i)}(\beta, \beta^0) h_t^{(l)}(\beta) \leq \frac{1}{T} \sum_{t=1}^T \varrho(\hat{\beta}, \beta^0) \bar{s}_t \bar{h}_t.$$

Ahora, usando la desigualdad de Hölder,

$$\begin{aligned} E |\bar{s}_t \bar{h}_t| &\leq \|\bar{s}_t\|_2 \|\bar{h}_t\|_2 \\ &\leq K_{s,h} < \infty, \end{aligned}$$

obtenemos como antes que, $\frac{1}{T} \sum_{t=1}^T \bar{s}_t \bar{h}_t = O_p(1)$, $\varrho(\hat{\beta}, \beta^0) = o_p(1)$ y por lo tanto

$$\lim_{T \rightarrow \infty} P \left[\sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T s_t^{(i)}(\beta, \beta^0) h_t^{(l)}(\beta) > \eta \right] \xrightarrow{T \rightarrow \infty} 0,$$

$$\lim_{T \rightarrow \infty} P \left[\sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T s_t^{(i)}(\beta, \beta^0) h_t^{(l)}(\beta^0) > \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Para el tercer sumando, $\frac{1}{T} \sum_{t=1}^T v_t(\beta, \beta^0) h_t^{i,l}(\beta)$, se puede probar, usando la definición (2.16) y $\phi_5 = \phi_5^0$, que $\bar{v}_t(\beta^0, \varrho(\beta, \beta^0)) = \varrho(\beta, \beta^0) \bar{v}_t$ donde \bar{v}_t no depende de ϕ y $E(\bar{v}_t^4) \leq K_s < \infty$. Por la desigualdad de Hölder y (2.53)

$$E \left[\bar{v}_t \sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} h_t^{i,l}(\beta) \right] \leq K_{v,h} < \infty.$$

Así, como en el caso anterior, $\frac{1}{T} \sum_{t=1}^T \bar{v}_t \bar{h}_t^{i,l} = O_p(1)$, $\varrho(\hat{\beta}, \beta^0) = o_p(1)$ y

$$\lim_{T \rightarrow \infty} P \left[\sup_{\beta^* / \varrho(\beta^*, \beta^0) < \varrho(\hat{\beta}, \beta^0)} \frac{1}{T} \sum_{t=1}^T v_t(\beta, \beta^0) h_t^{i,l}(\beta) > \eta \right] = 0.$$

Lo cual, con los resultados anteriores, prueba el lema. ■

Antes de presentar el último lema necesario para la prueba del Teorema 2.3, definimos

$$D_T(\beta^0) = \frac{1}{T^{1/2}} \sum_{t=1}^T \varepsilon_t h_t(\beta^0) = \frac{1}{T^{1/2}} \sum_{t=1}^T \varepsilon_t \left(h_t^{(1)}(\beta^0), h_t^{(2)}(\beta^0), h_t^{(3)}(\beta^0), h_t^{(4)}(\beta^0) \right)'$$

$$\Omega = E[\varepsilon_t^2 h_t(\beta^0) h_t'(\beta^0)] = \sigma_\varepsilon^2 E[h_t(\beta^0) h_t'(\beta^0)] = \sigma_\varepsilon^2 H(\beta^0),$$

usando en la última igualdad que ε_t es *iid* y $h_t(\beta^0)$ \mathfrak{F}_{t-1} -medible. El siguiente lema describe el comportamiento asintótico de $D_T(\beta^0)$.

Lema 2.8 *Bajo los supuestos G.0, G.2, A.1 y A.2, $D_T(\beta^0) \xrightarrow{d} N(\vec{0}, \Omega)$.*

Prueba del Lema 2.8. Si definimos

$$E(\varepsilon_t^2 h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0)) = \sigma_\varepsilon^2 \sigma_{h,i,l},$$

se tiene que

$$\varepsilon_t^2 h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0) - \sigma_\varepsilon^2 \sigma_{h,i,l}$$

es L_2 - *mixingala* para $i, l \leq 4$. Usando el supuesto A.2 definimos

$$\sigma_{\varepsilon hil}^4 = E \left(\varepsilon_t^2 h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0) - \sigma_\varepsilon^2 \sigma_{h,i,l} \right)^2 < \infty. \quad (2.54)$$

Por lo tanto usando una LLN para L_2 - *mixingalas*, obtenemos que

$$\Omega_T(i, l) = \frac{1}{T} \sum_{t=1}^T \varepsilon_t^2 h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0) \xrightarrow{as} \Omega(i, l) = \sigma_\varepsilon^2 \sigma_{h,i,l} \quad (2.55)$$

Por otro lado, definimos $X_{t,T} = \sum_{i=1}^4 \mu_i \frac{\varepsilon_t h_t^{(i)}(\beta^0)}{\sqrt{T}}$, con $\|\mu\| = 1$. Entonces usando (2.55),

$$\sum_{t=1}^T X_{t,T}^2 = \sum_{t=1}^T \sum_{i=1}^4 \mu_i^2 \left(\frac{\varepsilon_t h_t^{(i)}(\beta^0)}{\sqrt{T}} \right)^2 + 2 \sum_{t=1}^T \sum_{i=2}^4 \sum_{l=1}^{i-1} \mu_i \mu_l \frac{\varepsilon_t^2 h_t^{(i)}(\beta^0) h_t^{(l)}(\beta^0)}{T} \xrightarrow{as} \sigma_{x,\mu}^2 < \infty. \quad (2.56)$$

Además, teniendo en cuenta que $\varepsilon_t h_t^{(i)}(\beta^0)$ es una *mds*, $X_{t,T}^* = X_{t,T} / \sigma_{x,\mu}$ también será una *mds*. Por lo tanto, para probar que $\sum_{t=1}^T X_{t,T}^2$ converge a una $N(0, 1)$, usaremos un Teorema Central del Límite (CLT) para martingalas, por ejemplo, el propuesto por Davidson, Teorema 24.3, para el que debemos probar que

a) $\sum_{t=1}^T X_{t,T}^{*2} \xrightarrow{as} 1.$

b) $\max_{1 \leq t \leq T} |X_{t,T}^*| \xrightarrow{pr} 0.$

La condición a) se obtiene de (2.56). La condición b) se cumple dado que $\varepsilon_t h_t^{(i)}(\beta^0)$ es uniformemente integrable, lo que se sigue de la ecuación (2.54), (ver los Teoremas 24.4 y 12.10 de Davidson). Por lo tanto, con estas dos condiciones y el Teorema 24.3 de Davidson, se obtiene que $\sum_{t=1}^T X_{t,T}^{*2} \xrightarrow{d} N(0, 1)$. Dado que este resultado es válido para todo μ tal que $\|\mu\| = 1$, se tiene que $D_T(\beta^0) \xrightarrow{d} N(\vec{0}, \Omega)$, y con ello la prueba del lema. ■

Ahora ya estamos en disposición de probar el Teorema 2.3.

Prueba del Teorema 2.3. Usando los lemas 2.7 y 2.6, obtenemos que

$$Q_T(\hat{\beta}) - Q_T(\beta^0) = T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \beta^0) + 2T^{1/2}(\hat{\beta} - \beta^0)' D_T(\beta^0) + o_p(1),$$

con $H_T(\beta^0) = \frac{1}{T} \sum h'(\beta^0) h(\beta^0)$.

Define $T^{1/2}(\tilde{\beta} - \beta^0) = -H_T^{-1}(\beta^0) D_T(\beta^0)$, tal que

$$Q_T(\tilde{\beta}) - Q_T(\beta^0) = T(\tilde{\beta} - \beta^0)' H_T(\beta^{**}) (\tilde{\beta} - \beta^0) + 2T^{1/2}(\tilde{\beta} - \beta^0)' D_T(\beta^0) + o_p(1),$$

para $\beta^{**} \in (\tilde{\beta}, \beta^0)$. Entonces

$$Q_T(\hat{\beta}) - Q_T(\beta^0) = T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \beta^0) - 2T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\tilde{\beta} - \beta^0) + o_p(1)$$

$$Q_T(\tilde{\beta}) - Q_T(\beta^0) = -T(\tilde{\beta} - \beta^0)' H_T(\beta^0) (\tilde{\beta} - \beta^0) + o_p(1).$$

La última igualdad se obtiene dado que $-H_T^{-1}(\beta^0) D_T(\beta^0) = T^{1/2}(\tilde{\beta} - \beta^0)$ es $O_p(1)$, por lo tanto $\varrho(\tilde{\beta}, \beta^0) \xrightarrow{p} 0$ y $\varrho(\beta^{**}, \beta^0) \xrightarrow{p} 0$. Con esto, y siguiendo los pasos del lema 2.7, tenemos que $H_T(\beta^{**}) \xrightarrow{p} H_T(\beta^0)$. Ahora, dado que $Q_T(\tilde{\beta}) - Q_T(\beta^0) - [Q_T(\hat{\beta}) - Q_T(\beta^0)] \geq 0$,

$$\begin{aligned} 0 \leq & -T(\tilde{\beta} - \beta^0)' H_T(\beta^0) (\tilde{\beta} - \beta^0) - T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \beta^0) + \\ & + 2T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\tilde{\beta} - \beta^0) + o_p(1), \end{aligned}$$

y reasignando términos,

$$T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) + T(\tilde{\beta} - \hat{\beta})' H_T(\beta^0) (\tilde{\beta} - \beta^0) \leq o_p(1)$$

$$T(\hat{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) - T(\tilde{\beta} - \beta^0)' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) \leq o_p(1)$$

$$T(\hat{\beta} - \tilde{\beta})' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) \leq o_p(1).$$

Del lema 2.7, tenemos que $H_T(\beta^0) \xrightarrow{p} H(\beta^0) > 0$, con lo que

$$T(\hat{\beta} - \tilde{\beta})' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) \leq o_p(1) \Rightarrow T(\hat{\beta} - \tilde{\beta})' H_T(\beta^0) (\hat{\beta} - \tilde{\beta}) \leq o_p(1) \Rightarrow$$

$$T \|\hat{\beta} - \tilde{\beta}\|^2 C \leq o_p(1).$$

Finalmente, usando el lema 2.8, obtenemos la normalidad asintótica de $T^{1/2} (\tilde{\beta} - \beta^0)$,

$$T^{1/2} (\tilde{\beta} - \beta^0) \xrightarrow{d} N \left(\vec{0}, H^{-1} (\beta^0) \Omega H^{-1} (\beta^0) \right) \equiv N \left(\vec{0}, \sigma_\varepsilon^2 H^{-1} (\beta^0) \right),$$

que a su vez implica, a través del Teorema de Slutsky y dado que $\|\hat{\beta} - \tilde{\beta}\|^2 \leq o_p(1)$, que

$$T^{1/2} (\hat{\beta} - \beta^0) \xrightarrow{d} N \left(\vec{0}, \sigma_\varepsilon^2 H^{-1} (\beta^0) \right).$$

Esto finaliza la prueba del Teorema 2.3. ■

Prueba del Teorema 2.4. La prueba se obtiene directamente del lema 2.7, y teniendo en cuenta que,

$$\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2 = \frac{1}{T} \left(Q_T(\hat{\beta}) - Q_T(\beta^0) \right) \xrightarrow{p} 0.$$

■

Normalidad asintótica cuando r^0 es desconocido.

Durante esta sección mantenemos la notación $\beta = (\beta_1, \beta_2, \beta_3, \beta_4) = (\theta_1, \theta_2, \rho, \mu)$, con $\phi = (\beta, r)$ y $\phi^0 = (\beta^0, r^0)$. Sin embargo la dependencia de los procesos en el parámetro r aparecerá en la notación de forma explícita. Para un r dado, la función objetivo, como vimos en la prueba de consistencia, es derivable respecto a β_l para $l \leq 4$. Similar al Teorema 2.3, utilizando un desarrollo de Taylor de orden dos entorno a ϕ^0 ,

$$\begin{aligned} Q_T(\hat{\beta}, \hat{r}) - Q_T(\phi^0) &= \sum_{i=1}^4 \sum_{t=1}^T 2 \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t h_t^{(i)}(\beta^0, \hat{r}) + \\ &\sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t h_t^{il}(\beta^*, \hat{r}) \left(\hat{\beta}_i - \beta_i^0 \right) + 2 \sum_{i=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) v_t(\beta^0, \hat{r}, \phi^0) h_t^{(i)}(\beta^0, \hat{r}) + \\ &+ \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) v_t(\beta^*, \hat{r}, \phi^0) h_t^{i,l}(\beta^*, \hat{r}) \left(\hat{\beta}_i - \beta_i^0 \right) + \\ &+ \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) h_t^{(l)}(\beta^*, \hat{r}) h_t^{(i)}(\beta^*, \hat{r}) \left(\hat{\beta}_l - \beta_l^0 \right), \end{aligned} \quad (2.57)$$

donde $v_t(\beta, \hat{r}, \phi^0) = e_t(\beta, \hat{r}) - \varepsilon_t$, $h_t^{i,l}$ se definió en la sección anterior y β^* cumple que $\varrho(\beta^*, \beta^0) \leq \varrho(\hat{\beta}, \beta^0)$. La prueba del Teorema 2.5, como en el caso del Teorema 2.3, se seguirá de la siguiente igualdad:

$$\begin{aligned} Q_T(\hat{\beta}, \hat{r}) - Q_T(\phi^0) &= \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\hat{\beta}_i - \beta_i^0) H_t^{i,l}(\phi^0) (\hat{\beta}_i - \beta_i^0) + \\ &+ \sum_{i=1}^4 \sum_{t=1}^T 2 (\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\phi^0) + o_p(1). \end{aligned} \quad (2.58)$$

Para probar esta igualdad, usaremos los siguientes tres lemas.

Lema 2.9 *Bajo los supuestos del Teorema 2.5, para todo $\eta > 0$,*

$$\lim_{\nu \rightarrow 0} P \left[\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T [h_t^{(l)}(\phi) - h_t^{(l)}(\phi^0)] h_t^{(i)}(\phi) > \eta \right] = 0$$

para $l, i \leq 4$.

Prueba del Lema 2.9. Usando las definiciones (2.12) y (2.18),

$$\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T [h_t^{(l)}(\phi) - h_t^{(l)}(\phi^0)] h_t^{(i)}(\phi) \leq \frac{1}{T} \sum_{t=1}^T \bar{s}_t(\phi^0, \nu) \bar{h}_t^{(i)}.$$

Del lema 2.3 y las desigualdades de Markov y Hölder,

$$P \left[\frac{1}{T} \sum_{t=1}^T \bar{s}_t(\phi^0, \nu) \bar{h}_t^{(i)} > \eta \right] \leq \frac{\nu K_{sh}}{\eta},$$

que convergerá a 0 si ν lo hace, dado que $K_{sh} < \infty$.

■

Lema 2.10 *Bajo los supuestos del Teorema 2.5 y para $3/4 < \alpha < 1$,*

$$\begin{aligned} \lim_{T \rightarrow 0} P \left[\sup_{\varrho(r, r^0) < \frac{\Delta}{T^\alpha}} \frac{1}{T^{1/2}} \sum_{t=1}^T v_t(\beta^0, r, \phi^0) h_t^{(i)}(\beta^0, r) > \eta \right] &= 0 \\ \lim_{T \rightarrow 0} P \left[\sup_{\varrho(r, r^0) < \frac{\Delta}{T^\alpha}} \frac{1}{T^{1/2}} \sum_{t=1}^T \varepsilon_t [h_t^{(i)}(\beta^0, \hat{r}) - h_t^{(i)}(\phi^0)] > \eta \right] &= 0, \end{aligned}$$

para $i \leq 4$.

Prueba del Lema 2.10. De la definición de $v_t(\phi, \phi^0)$,

$$\sup_{\varrho(r, r^0) < \frac{\Delta}{T^\alpha}} v_t(\beta^0, r, \phi^0) h_t^{(i)}(\beta^0, r) \leq \bar{v}_t\left(\phi^0, \frac{\Delta}{T^\alpha}\right) \bar{h}_t^{(i)}.$$

Por la desigualdad de Hölder,

$$E \left| \bar{v}_t\left(\phi^0, \frac{\Delta}{T^\alpha}\right) \bar{h}_t^{(i)} \right| \leq \left\| \bar{v}_t\left(\phi^0, \frac{\Delta}{T^\alpha}\right) \right\|_{3/2} \left\| \bar{h}_t^{(i)} \right\|_3 \leq \frac{\Delta}{T^{2\alpha/3}} K_{vh}.$$

Por lo tanto,

$$P \left[\sup_{\varrho(r, r^0) < \frac{\Delta}{T^\alpha}} \frac{1}{T^{1/2}} \sum_{t=1}^T v_t(\beta^0, r, \phi^0) h_t^{(i)}(\beta^0, r) > \eta \right] \leq T^{\frac{1}{2} - \frac{2\alpha}{3}} K_{vh} \xrightarrow{T \rightarrow \infty} 0,$$

dado que $\alpha > 3/4$. Para el segundo límite, obtenemos las siguiente desigualdades,

$$\begin{aligned} \sup_{\varrho(r, r^0) < \frac{\Delta}{T^\alpha}} \left| h_t^{(i)}(\beta^0, \hat{r}) - h_t^{(i)}(\phi^0) \right| &\leq \bar{s}_t\left(\phi^0, \frac{\Delta}{T^\alpha}\right), \\ E \left[|\varepsilon_t| \bar{s}_t\left(\phi^0, \frac{\Delta}{T^\alpha}\right) \right] &\leq \frac{\Delta}{T^\alpha} K_{s\varepsilon}. \end{aligned}$$

Con esta última desigualdad, y siguiendo los mismos pasos que el primer caso, probamos el segundo límite del lema. ■

Lema 2.11 *Bajo los supuestos del Teorema 2.5,*

$$\begin{aligned} \lim_{\nu \rightarrow 0} P \left[\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T \varepsilon_t [h_t^{il}(\phi) - h_t^{il}(\phi^0)] > \eta \right] &= 0 \\ \lim_{\nu \rightarrow 0} P \left[\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T v_t(\phi, \phi^0) h_t^{i,l}(\phi) > \eta \right] &= 0, \end{aligned}$$

para $l, i \leq 4$.

Prueba del Lema 2.11. Para el primer límite, de la definición de $h_t^{il}(\phi)$ y siguiendo los mismos pasos que en el lema 2.3, se prueba que

$$E \left(\sup_{\varrho(\phi, \phi^*) < \nu} |h_t^{il}(\phi) - h_t^{il}(\phi^*)| \right) \leq \nu K_{hh},$$

con $K_{hh} < \infty$. De nuevo, usando la desigualdad de Markov y Hölder

$$P \left[\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T \varepsilon_t [h_t^{il}(\phi) - h_t^{il}(\phi^0)] > \eta \right] \leq \frac{\nu K_{hh} \sigma_\varepsilon}{\eta}.$$

Para el segundo límite, usando que $\left\| \sup_{\phi \in \Theta} h_t^{i,l}(\phi) \right\| < \infty$, se obtiene a través de la desigualdad de Hölder que,

$$E \left(\sup_{\varrho(\phi, \phi^*) < \nu} \left| v_t(\phi^0, \nu) h_t^{i,l}(\phi) \right| \right) \leq \nu^{2/3} K_{vh},$$

por lo tanto,

$$P \left[\sup_{\varrho(\phi, \phi^0) < \nu} \frac{1}{T} \sum_{t=1}^T v_t(\phi, \phi^0) h_t^{i,l}(\phi) > \eta \right] \leq \frac{\nu^{2/3} K_{vh}}{\eta}.$$

Ambas probabilidades convergerán a 0 cuando lo haga ν . ■

Prueba del Teorema 2.5 . Para probar la igualdad (2.58) empezamos probando

que

$$\sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) \left[h_t^{(l)}(\beta^*, \hat{r}) h_t^{(i)}(\beta^*, \hat{r}) - h_t^{(l)}(\phi^0) h_t^{(i)}(\phi^0) \right] \left(\hat{\beta}_l - \beta_l^0 \right) = o_p(1).$$

Para ello es suficiente probar que $T^{-1} \sum_{t=1}^T \left[h_t^{(l)}(\beta^*, \hat{r}) h_t^{(i)}(\beta^*, \hat{r}) - h_t^{(l)}(\phi^0) h_t^{(i)}(\phi^0) \right] = o_p(1)$ para todo $i, l \leq 4$. Este resultado se obtiene de la aplicación del lema 2.9 con $\nu_T = T^{-d}$ con $0 < d < 1/2$ y el Teorema 2.2 que garantiza que el $\lim_{T \rightarrow \infty} P[\varrho(\phi, \phi^0) < \nu_T] = 1$.

Con un argumento similar, usando el lema 2.10 y dado que el $\lim_{T \rightarrow \infty} P[|\hat{r} - r^0| < \frac{\Delta}{T^\alpha}] = 1$ para $3/4 < \alpha < 1$,

$$\begin{aligned} \sum_{i=1}^4 \sum_{t=1}^T 2 \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t \left[h_t^{(i)}(\beta^0, \hat{r}) - h_t^{(i)}(\phi^0) \right] &= o_p(1) \\ \sum_{i=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) v_t(\beta^0, \hat{r}, \phi^0) h_t^{(i)}(\beta^0, \hat{r}) &= o_p(1). \end{aligned}$$

Usando los resultados del lema 2.11 y con la argumentación anterior, obtenemos que

$$\begin{aligned} \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T 2 \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t \left[h_t^{il}(\beta^*, \hat{r}) - h_t^{il}(\phi^0) \right] \left(\hat{\beta}_l - \beta_l^0 \right) &= o_p(1) \\ \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) v_t(\beta^*, \hat{r}, \phi^0) h_t^{i,l}(\beta^*, \hat{r}) \left(\hat{\beta}_l - \beta_l^0 \right) &= o_p(1). \end{aligned}$$

Finalmente, del lema 2.6, tenemos el siguiente resultado,

$$\sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T 2 \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t h_t^{il}(\phi^0) \left(\hat{\beta}_i - \beta_i^0 \right) = o_p(1).$$

Así, con estos resultados es directo probar la igualdad de la ecuación (2.58). La normalidad asintótica de $T^{1/2} \left(\hat{\beta} - \beta^0 \right)$ a partir de la igualdad (2.58) esta desarrollada en el Teorema 2.3 y la consistencia de la matriz de varianzas y covarianzas se obtiene fácilmente a partir del lema 2.9. ■

Contraste para la existencia de Threshold cuando r^0 es desconocido.

El esquema de esta sección es el siguiente. En primer lugar probaremos que el estadístico del Teorema 2.6 puede ser escrito como una función del siguiente proceso:

$$H_T^{-1}(r) s_T(r) + o_p(1),$$

donde el $o_p(1)$ es uniforme sobre r , s_T es una *mds* y $H_T(r)$ converge uniformemente en r a $H(r)$. Para ello, necesitamos, en primer lugar, que la tasa de convergencia obtenida en el Teorema 2.2 para todos los parámetros, excepto el parámetro threshold r , se cumpla de forma uniforme para todo r y bajo la hipótesis nula de no efecto threshold. Los pasos para obtener este resultado son similares a los seguidos en el Teorema 2.2, aunque en este caso asumiremos que los estimadores son consistentes, lo cual se obtiene fácilmente de dicho Teorema. En segundo lugar, probaremos que el proceso empírico $s_T(r)$ converge en distribución uniformemente a $s(r)$, esto es, debemos probar la convergencia finito dimensional, $(s_T(r_1), s_T(r_2), \dots, s_T(r_k)) \xrightarrow{d} (s(r_1), s(r_2), \dots, s(r_k))$, y que el proceso es tight. Al final de la sección se encuentra la prueba del Teorema 2.6, que reagrupa los resultados obtenidos hasta entonces.

Para esta sección, denotamos $\beta = (\beta_1, \beta_2, \beta_3, \beta_4) = (\theta, \alpha, \rho, \mu)$, con $\phi = (\beta, r)$, $\phi^0 =$

(β^0, r^0) y

$$e_t(\phi) = x_t - \rho x_{t-1} - \mu + \theta e_{t-1}(\phi) + \alpha e_{t-1}(\phi) 1(z_{t-1} < r)$$

$$v_t(\phi, \phi^0) = (\rho^0 - \rho) x_{t-1} + (\mu^0 - \mu) + (\theta - \theta^0) \varepsilon_{t-1} + \alpha e_{t-1}(\phi) 1(z_{t-1} < r) + \theta v_{t-1}(\phi, \phi^0).$$

Empezamos por la tasa uniforme de convergencia, esto es, probando que el $\sup_{r \in [\underline{r}, \bar{r}]} \sqrt{T} \left(\varrho \left(\hat{\beta}(r), \beta^0 \right) \right) = O_p(1)$. Para obtener este resultado redefinimos $B_{\Delta, \epsilon} = \left\{ \phi \in \Theta : \sqrt{\frac{\Delta}{T}} \leq |\phi_l - \phi_l^0| = \delta_{\phi, l}^0 \leq \epsilon \text{ para } l = 1, \dots, 4 \text{ y } r \in [\underline{r}, \bar{r}] \right\}$. Con esta nueva definición de $B_{\Delta, \epsilon}$, $e_t(\phi)$ y $v_t(\phi, \phi^0)$ la tasa de convergencia uniforme propuesta se obtendrá de las siguientes dos proposiciones.

Proposición 2.19 Para todo $\eta, \varepsilon > 0 \exists \Delta < \infty$ tal que

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t(v_t(\phi, \phi^0))}{T \varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon.$$

Proposición 2.20 Para todo $\varepsilon > 0 \exists \Delta < \infty$ y $\eta > 0$ tal que

$$P \left[\inf_{\phi \in B_{\Delta, \epsilon}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho^2(\beta, \beta^0)} > \eta \right] \geq 1 - \varepsilon.$$

Prueba de la Proposición 2.19. Dado que para un r fijo, $Q_T(\beta, r)$ es diferenciable respecto a β_i e $i = 1, \dots, 4$, tenemos que

$$\sum_{t=1}^T \varepsilon_t(v_t(\phi, \phi^0)) = \sum_{t=1}^T \sum_{i=1}^4 (\beta_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\beta^0, r) + \sum_{t=1}^T \sum_{i, l=1}^4 (\beta_i - \beta_i^0) \varepsilon_t h_t^{i, l}(\beta^*, r) (\beta_l - \beta_l^0), \quad (2.59)$$

con $\varrho(\beta, \beta^*) \leq \varrho(\beta, \beta^0)$ y

$$h_t^{(1)}(\beta, r) = e_{t-1}(\beta, r) + \theta_{t-1} h_t^{(1)}(\beta, r)$$

$$h_t^{(2)}(\beta, r) = e_{t-1}(\beta, r) 1(z_{t-1} < r) + \theta_{t-1} h_{t-1}^{(2)}(\beta, r)$$

$$h_t^{(3)}(\beta, r) = -x_{t-1} + \theta_{t-1} h_{t-1}^{(3)}(\beta, r)$$

$$h_t^{(4)}(\beta, r) = -1 + \theta_{t-1} h_{t-1}^{(4)}(\beta, r), \quad (2.60)$$

con $\theta_{t-1} = \theta 1(z_{t-1} > r) + (\theta + \alpha) 1(z_{t-1} < r)$ y $e_t(\beta^0, r) = \varepsilon_t$. Dado que la definición de $e_t(\beta, r)$ ha cambiado, también cambiarán sus derivadas. Para el primer sumando de la ecuación (2.59), tenemos que

$$\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T \sum_{i=1}^4 (\beta_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\beta^0, r)}{T \varrho^2(\beta, \beta^0)} \right| \leq \sum_{i=1}^4 \sup_{r \in [\underline{r}, \bar{r}]} \left| \frac{\sum_{t=1}^T \varepsilon_t h_t^{(i)}(\beta^0, r)}{T^{1/2} \Delta} \right|.$$

De las definiciones de (2.60), solo $h_t^{(2)}(\beta^0, r)$ dependerá de r . Así, para $i = 1, 3, 4$ $T^{-1/2} \sum_{t=1}^T \varepsilon_t h_{t-1}^{(i)}(\beta^0, r) = O_p(1)$. Como probaremos más adelante, $T^{-1/2} \sum_{t=1}^T \varepsilon_t h_{t-1}^{(2)}(\beta^0, r)$ es tight en r , así que $\sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \sum_{t=1}^T \varepsilon_t h_{t-1}^{(2)}(\beta^0, r) = O_p(1)$. Por lo tanto, tomando Δ suficientemente grande, para $\forall \eta > 0$ y $\varepsilon > 0$,

$$P \left[\sum_{i=1}^4 \sup_{r \in [\underline{r}, \bar{r}]} \left| \frac{\sum_{t=1}^T \varepsilon_t h_{t-1}^{(i)}(\beta^0, r)}{T^{1/2} \Delta} \right| > \eta \right] \leq \varepsilon.$$

Para el segundo sumando,

$$\sup_{\phi \in B_{\Delta, \epsilon}} \frac{\sum_{t=1}^T \sum_{i,l=1}^4 (\beta_i - \beta_i^0) \varepsilon_t h_t^{i,l}(\beta^*, r) (\beta_l - \beta_l^0)}{T \varrho^2(\beta, \beta^0)} \leq \sum_{i,l=1}^4 \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t h_t^{i,l}(\phi)}{T} \right|.$$

Como en otras ocasiones, para un ϕ dado, $T^{-1} \left| \sum_{t=1}^T \varepsilon_t h_t^{i,l}(\phi) \right| = o_p(1)$, luego necesitamos una ULLN para completar la prueba de esta proposición. La prueba es la misma que para el sumando (2.24), dado que $\left| h_{t-1}^{i,l}(\phi) - h_{t-1}^{i,l}(\phi^*) \right| \leq \varrho(\phi, \phi^0) \bar{h}_{t-1}^{i,l}$ con $\bar{h}_{t-1}^{i,l}$ L_2 -mixingala de tamaño $-1/2$. Este resultado se obtiene siguiendo los mismos pasos que para $h_{t-1}^{(i)}(\phi)$ en los lemas 2.3 y 2.4. Con esta ley uniforme, y el resultado para el primer sumando de (2.59) se completa la prueba de la proposición. ■

Prueba de la Proposición 2.20. Dada la nueva definición de $v_t(\phi, \phi^0)$, tenemos que

$$\inf_{\phi \in B_{\Delta, \epsilon}} \sum_{t=1}^T v_t^2(\phi, \phi^0) = \inf_{\phi \in B_{\Delta, \epsilon}} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) \quad \text{para } k(\iota) = 2(\iota - 1) + 1,$$

con

$$\begin{aligned} W_{k(\iota)}(\phi, \phi^0) &\geq 1(z_{k(\iota)-1} > \bar{r}) (\delta_{\phi,3}^0 x_{t-1} + \delta_{\phi,4}^0 + \delta_{\phi,1}^0 \varepsilon_{t-1})^2 + \\ &\quad + 1(z_{k(\iota)-1} < \underline{r}) (\delta_{\phi,3}^0 x_{t-1} + \delta_{\phi,4}^0 + \delta_{\phi,2}^0 \varepsilon_{t-1})^2, \end{aligned}$$

donde en este caso, $\delta_{\phi,1}^0 = (\beta_1 - \beta_1^0)$ y $\delta_{\phi,2}^0 = \delta_{\phi,1}^0 + \beta_2$. Como en el caso de la proposición 2.18,

$$\inf_{\{\phi \in B_{\Delta,\epsilon}\}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho^2(\phi, \phi^0)} \geq \inf_c \frac{\sum_{t=1}^T c' M_{k(\iota)-1}(\bar{r}) c}{T} + \inf_c \frac{\sum_{t=1}^T c' M_{k(\iota)-1}(\underline{r}) c}{T}, \quad (2.61)$$

con $\max_i c_i^2 = 1$. Bajo los supuestos G.2 y A.3, como vimos en la proposición 2.18, cada sumando de la ecuación (2.61) convergerá al $\inf_c E [c' M_{k(\iota)-1}(\bar{r}) c]$ e $\inf_c E [c' M_{k(\iota)-1}(\underline{r}) c]$ respectivamente. En este caso, dividimos $B_{\Delta,\epsilon}$ en dos subconjuntos, el primero, $2|\beta_1 - \beta_1^0| > \beta_2$, y su complementario. Para el primero, $\inf_c E [c' M_{k(\iota)-1}(\bar{r}) c] > a_2^* > 0$ y para el segundo $\inf_c E [c' M_{k(\iota)-1}(\bar{r}) c] > a_3^* > 0$. Por lo tanto, como en la proposición 2.18, obtenemos que

$$P \left[\inf_{\{\phi \in B_{\Delta,\epsilon}\}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho^2(\phi, \phi^0)} > \eta \right] \geq 1 - \varepsilon.$$

■

Con la prueba de ambas proposiciones, queda probado que $\sup_{r \in [\underline{r}, \bar{r}]} \sqrt{T} \left(\varrho(\hat{\beta}(r), \beta^0) \right) = O_p(1)$. Ahora pasamos a probar que $T^{1/2}(\hat{\beta} - \beta^0) = H_T(r) s_T(r) + o_p(1)$. Usando un desarrollo de Taylor de orden dos para $Q_T(\hat{\beta}, r)$ en torno a $Q_T(\beta^0, r)$, tenemos que

$$\begin{aligned} Q_T(\hat{\beta}, r) - Q_T(\beta^0, r) &= \sum_{i=1}^4 \sum_{t=1}^T 2(\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\beta^0, r) + \\ &+ \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{i,l}(\beta^*, r) (\hat{\beta}_l - \beta_l^0) \\ &+ \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\hat{\beta}_i - \beta_i^0) H_{t-1}^{i,l}(\beta^*, r) (\hat{\beta}_l - \beta_l^0), \end{aligned}$$

con

$$H_{t-1}^{i,l}(\beta^*, r) = h_{t-1}^{(l)}(\beta^*, r) h_{t-1}^{(i)}(\beta^*, r),$$

y β^* tal que $\varrho(\beta^*, \beta^0) \leq \varrho(\hat{\beta}, \beta^0)$. El segundo sumando, $\varepsilon_t h_t^{i,l}(\beta^*, r)$, como vimos en la proposición 2.19 convergerá uniformemente a 0. Para el tercer sumando, utilizamos los siguientes dos lemas.

Lema 2.12 *Bajo los supuestos del Teorema 2.6*

$$\sup_{r \in [\underline{r}, \bar{r}]} \sup_{\beta / \varrho(\beta, \beta^0) \leq \sqrt{\frac{\Delta}{T}}} \frac{1}{T} \sum_{t=1}^T H_{t-1}^{i,l}(\beta, r) - H_{t-1}^{i,l}(\beta^0, r) = o_p(1).$$

Prueba del Lema 2.12. De la definición de $H_{t-1}^{i,l}(\phi, r)$, obtenemos que

$$\begin{aligned} H_{t-1}^{i,l}(\beta, r) - H_{t-1}^{i,l}(\beta^0, r) &= \left[h_{t-1}^{(l)}(\beta^0, r) - h_{t-1}^{(l)}(\beta, r) \right] h_{t-1}^{(i)}(\beta^0, r) + \\ &\quad + \left[h_{t-1}^{(i)}(\beta^0, r) - h_{t-1}^{(i)}(\beta, r) \right] h_{t-1}^{(l)}(\beta, r) \\ &\leq 2\varrho(\hat{\beta}, \beta^0) \bar{s}_{t-1} \bar{h}_{t-1} \end{aligned}$$

De las definiciones en (2.60), para un r dado, tenemos que $\left| h_{t-1}^{(l)}(\beta^0, r) - h_{t-1}^{(l)}(\beta, r) \right| \leq \varrho(\beta, \beta^0) \bar{s}_{t-1}$, donde \bar{s}_{t-1} no depende de ϕ . Dado que $E(\bar{s}_{t-1} \bar{h}_{t-1}) < K_{sh} < \infty$ y $\sup_{r \in [\underline{r}, \bar{r}]} \varrho(\hat{\beta}, \beta^0) = o_p(1)$, el resultado del lema se obtiene directamente. ■

Lema 2.13 *Bajo los supuestos del Teorema 2.6*

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{T} \sum_{t=1}^T H_{t-1}^{i,l}(\beta^0, r) - H^{i,l}(r) = o_p(1),$$

con $E(H_{t-1}^{i,l}(\beta^0, r)) = H^{i,l}(r)$.

Prueba del Lema 2.13. Por definición, ver ecuación (2.60),

$$\begin{bmatrix} h_t^{(1)}(\beta^0, r) \\ h_t^{(2)}(\beta^0, r) \\ h_t^{(3)}(\beta^0, r) \\ h_t^{(4)}(\beta^0, r) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^t \theta^{0j-1} \varepsilon_{t-j} \\ \sum_{j=1}^t \theta^{0j-1} 1(z_{t-j} < r) \varepsilon_{t-j} \\ \sum_{j=1}^t \theta^{0j-1} x_{t-1} \\ \sum_{j=1}^t \theta^{0j-1} \end{bmatrix}.$$

Como en los lemas 2.4 y 2.5, se puede probar que $h_t^{(l)}(\beta^0, r) h_t^{(i)}(\beta^0, r)$ para $l, i = 1, \dots, 4$ son L_2 -mixingales de tamaño $-1/2$. Por lo tanto, para $l, i = 1, 3, 4$ la prueba del lema

es directa, ya que $h_t^{(l)}(\beta^0, r) h_t^{(i)}(\beta^0, r)$ no depende de r . Para el resto de casos, se prueba que

$$\sup_{r/|r-r'| < \nu} h_t^{(l)}(\beta^0, r) h_t^{(i)}(\beta^0, r) - h_t^{(l)}(\beta^0, r^*) h_t^{(i)}(\beta^0, r^*) \leq \bar{s}_t(\beta^0, \nu) \bar{h}_t$$

$$E(\bar{s}_t(\beta^0, \nu) \bar{h}_t) \leq \nu^{1/2} K_{sh}$$

y $\bar{s}_t(\beta^0, \nu) \bar{h}_t$ es un proceso L_2 -mixingala, de tamaño $-1/2$, y constante $\|\bar{s}_t(\beta^0, \nu) \bar{h}_t\|_2$. Para ello, primero se prueba que $\bar{s}_t(\beta^0, \nu)$ y \bar{h}_t son L_2 -NED, de tamaño $-\infty$ como en el lema 2.4. Posteriormente, usando el Teorema 2.14 en el Anexo 1, se prueba que también lo es el producto, $\bar{s}_t(\beta^0, \nu) \bar{h}_t$. La propiedad L_2 -mixingala, se obtiene del Teorema 2.12 en el Anexo 1. Así, siguiendo los pasos de la prueba de la desigualdad (2.34) en la proposición 2.16, se obtiene el lema cuando l o i son iguales a 2. ■

Usando el lema 2.12, se tiene que,

$$\begin{aligned} Q_T(\hat{\beta}, r) - Q_T(\beta^0, r) &= \sum_{i=1}^4 \sum_{t=1}^T 2(\hat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{(i)}(\beta^0, r) + \\ &+ \frac{1}{2} \sum_{i=1}^4 \sum_{l=1}^4 (\hat{\beta}_i - \beta_i^0) H_T^{i,l}(\beta^0, r) (\hat{\beta}_l - \beta_l^0) + o_p(1) \end{aligned}$$

donde el $o_p(1)$ es uniforme en r . Como en la prueba del Teorema 2.3, $T^{1/2}(\hat{\beta}(r) - \beta^0)$ para un r dado se puede escribir como

$$\left[H_T^{i,l}(\beta^0, r) \right]^{-1} T^{-1/2} \sum_{t=1}^T \varepsilon_t h_t(\beta^0, r) + o_p(1),$$

donde el $o_p(1)$ es uniforme en r y $H_T^{i,l}(\beta^0, r) = T^{-1} \sum_{t=2}^T H_{t-1}^{i,l}(\beta^0, r)$. Además, por el lema 2.13, $H_T^{i,l}(\beta^0, r)$ converge uniformemente en r a $H^{i,l}(\beta^0, r)$. Definimos a continuación $s_T(r) = T^{-1/2} \sum_{t=1}^T \varepsilon_t h_t(\beta^0, r)$, que para un r fijo será una *mds* estacionaria y cuadrado-integrable. Por lo tanto, dado que bajo $H_0 : \alpha = 0$, $\hat{\alpha}_T(r) = T^{1/2} R'(\hat{\beta}(r) - \beta^0)$, con $R' = (0, 1, 0, 0)$, y queda probado que el estadístico es un funcional de $H_T^{-1}(r) s_T(r)$.

Lema 2.14 *Bajo los supuestos G.0, G.2 y A.2, el proceso $s_T(r) = T^{-1/2} \sum_{t=1}^T \varepsilon_t h_t(\beta^0, r)$ es tight.*

Prueba del Lema 2.14. Teniendo en cuenta que $h_t^{(i)}(\beta^0, r)$ para $i = 1, 3, 4$ no depende de r , la tightness del proceso $s_T(r)$ dependerá de la tightness del proceso $T^{-1/2} \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j}$. Por el Teorema 4.2 de Billingsley (1968), y usando la siguiente descomposición,

$$\begin{aligned} T^{-1/2} \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} &= T^{-1/2} \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{m-1} \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} + \\ &+ T^{-1/2} \sum_{t=m}^T \varepsilon_t \sum_{j=m}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} + \\ &+ T^{-1/2} \sum_{t=1}^m \varepsilon_t \sum_{j=1}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j}, \end{aligned}$$

será suficiente probar que para todo $\eta > 0$,

$$\begin{aligned} \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[\sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=1}^m \varepsilon_t \sum_{j=1}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} \right| > \eta/2 \right] &= 0 \\ \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[\sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=m}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} \right| > \eta/2 \right] &= 0, \end{aligned}$$

y la tightness del proceso $T^{-1/2} \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{m-1} \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j}$.

Claramente, para la primera probabilidad,

$$\lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[\sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=1}^m \varepsilon_t \sum_{j=1}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} \right| \geq \eta/2 \right] \leq \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{mK}{T^{-1/2}\eta^2} = 0.$$

Para la segunda probabilidad, usaremos la partición que se obtiene de $r_l = \underline{r} + T^{-1/2}l$, para $l = 0, \dots, L = (\bar{r} - \underline{r})T^{1/2}$, con la que obtenemos que

$$\sup_{r \in [\underline{r}, \bar{r}]} \sum_{t=m}^T \varepsilon_t \sum_{j=m}^t \theta^{0,j-1} 1(z_{t-j} < r) \varepsilon_{t-j} = \theta^{0,m-1} \sup_{r \in [\underline{r}, \bar{r}]} \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0,j-1} 1(z_{t-m-j} < r) \varepsilon_{t-m-j},$$

y

$$\begin{aligned} & \sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(z_{t-m-j} < r) \varepsilon_{t-m-j} \right| \leq \\ & \max_l T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(z_{t-m-j} < r_l) \varepsilon_{t-m-j} \right| + \\ & + \max_l \sup_{r \in [r_l, r_{l+1}]} T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(r_l < z_{t-m-j} < r) \varepsilon_{t-m-j} \right|. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} & P \left[|\theta^0|^{m-1} \sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(z_{t-m-j} < r) \varepsilon_{t-m-j} \right| > \eta \right] \leq \\ & P \left[|\theta^0|^{m-1} \max_l T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(z_{t-m-j} < r_l) \varepsilon_{t-m-j} \right| > \eta/2 \right] + \end{aligned} \quad (2.62)$$

$$+ P \left[|\theta^0|^{m-1} \max_l T^{-1/2} \sum_{t=m}^T |\varepsilon_t| \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(r_l < z_{t-m-j} < r_{l+1}) |\varepsilon_{t-m-j}| > \eta/2 \right]. \quad (2.63)$$

Para obtener los resultados deseados respecto a las probabilidades (2.62) y (2.63) definimos la siguiente mds $U_t(r) = \varepsilon_t \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(z_{t-m-j} < r) \varepsilon_{t-m-j}$. Bajo los supuestos del Teorema 2.6, $E(U_t^2) \leq K_\sigma < \infty$. Igualmente definimos $U_{1,t}(r_1, r) = \sum_{j=1}^{t-m} \theta^{0^{j-1}} 1(r_l < z_{t-m-j} < r) |\varepsilon_{t-m-j}|$, que bajo los supuestos del Teorema 2.6 cumplirá que, $U_{1,t}(r_1, r) - E(U_{1,t}(r_1, r))$ y $U_{1,t}^2(r_1, r) - E(U_{1,t}^2(r_1, r))$ son $L_2 - NED$ de tamaño $-\infty$ en (z_t, ε_t) con constante $|r_1 - r|^{1/2} K$. Dado que (z_t, ε_t) son $\phi - mixing$, usando el Teorema 2.12, se obtiene que son $L_2 - mixingalas$ de tamaño $-1/2$ con constante $|r_1 - r|^{1/2} K$. Por lo tanto,

$$\begin{aligned} & E(U_{1,t}^4(r_1, r)) \leq |r_1 - r| K_\sigma \\ & E \left(\sum_{t=n}^T U_{1,t}^2(r_1, r) - E(U_{1,t}^2(r_1, r)) \right)^2 = T |r_1 - r| K_\sigma. \end{aligned}$$

Para la primera probabilidad, ecuación (2.62), usamos el Teorema 12.2 de Billingsley (1968). Para ello, debemos probar que para cualquier entero $i < j$ entre 0 y L (definidos

en la partición de $[r, \bar{r}]$,

$$\begin{aligned} P \left[T^{-1/2} \sum_{t=m}^T U_t(r_i) - U_t(r_j) \geq \eta \right] &= P \left[T^{-1/2} \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0j-1} 1(r_i < z_{t-m-j} < r_j) \varepsilon_{t-m-j} \geq \eta \right] \\ &\leq |r_i - r_j|^\pi K, \end{aligned}$$

con $\pi > 1$. Para ello, usando la desigualdad de Burkholder, los resultados previos y que $(\varepsilon_t^2 - E(\varepsilon_t^2))^2 U_{1,t}^4(r_i, r_j)$ es una *mds*, tenemos que

$$\begin{aligned} E \left(\sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0j-1} 1(r_i < z_{t-m-j} < r_j) \varepsilon_{t-m-j} \right)^4 &\leq E \left(\sum_{t=m}^T \varepsilon_t^2 U_{1,t}^2(r_i, r_j) \right)^2 \\ E \left(\sum_{t=m}^T \varepsilon_t^2 U_{1,t}^2(r_i, r_j) \right)^2 &= E \left(\sum_{t=m}^T (\varepsilon_t^2 - E(\varepsilon_t^2)) U_{1,t}^2(r_i, r_j) + E(\varepsilon_t^2) U_{1,t}^2(r_i, r_j) \right)^2 \\ &\leq \sum_{t=m}^T (\varepsilon_t^2 - E(\varepsilon_t^2))^2 U_{1,t}^4(r_i, r_j) + E \left(\sum_{t=m}^T E(\varepsilon_t^2) U_{1,t}^2(r_i, r_j) \right)^2 \leq \\ &\leq T |r_j - r_i| K_\sigma + E \left(\sum_{t=m}^T E(\varepsilon_t^2) U_{1,t}^2(r_i, r_j) \right)^2. \end{aligned}$$

Con los resultados sobre los momentos de $U_{1,t}^2(r_i, r_j)$,

$$\begin{aligned} E \left(\sum_{t=m}^T U_{1,t}^2(r_i, r_j) - E(U_{1,t}^2(r_i, r_j)) + E(U_{1,t}^2(r_i, r_j)) \right)^2 &\leq TC |r_j - r_i| + \\ &\quad + E \left(\sum_{t=m}^T E(U_{1,t}^2(r_i, r_j)) \right)^2 \\ &\leq TC |r_j - r_i| + T^2 |r_j - r_i|^2 K_\sigma. \end{aligned}$$

Reasignando,

$$\begin{aligned} E \left(T^{-1/2} \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0j-1} 1(r_i < z_{t-m-j} < r_j) \varepsilon_{t-m-j} \right)^4 &\leq C \frac{|r_j - r_i|}{T} + |r_j - r_i|^2 K_\sigma \\ &\leq |r_j - r_i|^2 K_\sigma^*, \end{aligned}$$

usando que $|r_i - r_j| > T^{-1/2}$. Del Teorema 12.2 de Billingsley (1968),

$$\lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[|\theta^0|^{m-1} \max_l T^{-1/2} \left| \sum_{t=m}^T U_t(r_l) \varepsilon_{t-m-j} \right| > \eta/2 \right] \leq \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{|\theta^0|^{2m-2} K_{\eta, \sigma}}{\eta^4} = 0.$$

Para la segunda probabilidad, (2.63),

$$\begin{aligned} \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[\max_l \frac{|\theta^0|^{m-1}}{T^{1/2}} \left| \sum_{t=m}^T [|\varepsilon_t| - E(|\varepsilon_t|)] U_{1,t}(r_l, r_{l+1}) \right| > \eta/4 \right] &\leq \lim_{m \rightarrow \infty} \frac{|\theta^0|^{2m-2} K}{\eta^2} = 0 \\ \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[\max_l \frac{|\theta^0|^{m-1}}{T^{1/2}} \left| \sum_{t=m}^T [U_{1,t}(r_l, r_{l+1}) - E(U_{1,t}(r_l, r_{l+1}))] \right| > \eta/8 \right] &\leq \lim_{m \rightarrow \infty} \frac{|\theta^0|^{2m-2} K}{\eta^2} = 0 \\ \lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} \theta^{0m-1} \max_l T^{-1/2} \sum_{t=m}^T E(U_{1,t}(r_l, r_{l+1})) &< \lim_{m \rightarrow \infty} |\theta^0|^{2m-2} = 0, \end{aligned}$$

donde para la primera desigualdad se usa que $[|\varepsilon_t| - E(|\varepsilon_t|)] U_{1,t}(r_l, r_{l+1})$ es una *mds* con segundo momento acotado por $T^{-1/2}K$. Para la segunda, usamos que

$[U_{1,t}(r_l, r_{l+1}) - E(U_{1,t}(r_l, r_{l+1}))]$ es una L_2 -*mixingala* de tamaño $-1/2$ con constante $T^{-1/2}K$. Resumiendo,

$$\lim_{m \rightarrow \infty} \lim_{T \rightarrow \infty} P \left[|\theta^0|^{m-1} \sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \left| \sum_{t=m}^T \varepsilon_t \sum_{j=1}^{t-m} \theta^{0j-1} 1(z_{t-m-j} < r) \varepsilon_{t-m-j} \right| \geq \eta \right] = 0.$$

Por lo tanto, las condiciones del Teorema 4.2 de Billingsley (1968) son satisfechas y será suficiente probar la tightness de

$$P \left[\sup_{r \in [\underline{r}, \bar{r}]} T^{-1/2} \sum_{t=m}^T \varepsilon_t \theta^{0j-1} 1(z_{t-j} > r) \varepsilon_{t-j} \geq \eta \right] \leq \epsilon$$

para $j = 1, \dots, m-1$, dado que la suma finita de procesos tight será tight.

Así, del lema 3.1 de Koul y Stute (1999), para obtener la tightness del proceso $T^{-1/2} \sum_{t=m}^T \varepsilon_t \theta^{0j-1} 1(z_{t-j} < r) \varepsilon_{t-j}$, será suficiente probar que para todo $\underline{r} < r_1 < r_2 < r_3 < \bar{r}$,

$$E \left\{ [\alpha_T(r_1) - \alpha_T(r_2)]^2 [\alpha_T(r_2) - \alpha_T(r_3)]^2 \right\} \leq |r_1 - r_2| |r_2 - r_3| K,$$

donde,

$$\alpha_T(r) = T^{-1/2} \sum_{t=m}^T \varepsilon_t 1(z_{t-j} < r) \varepsilon_{t-j}.$$

Concretamente, de las definiciones anteriores tenemos que,

$$\begin{aligned} [\alpha_T(r_1) - \alpha_T(r_2)]^2 [\alpha_T(r_2) - \alpha_T(r_3)]^2 &= \\ T^{-2} \left[\sum_{t=n}^T \varepsilon_t 1(r_2 < z_{t-j} \leq r_1) \varepsilon_{t-j} \right]^2 & \left[\sum_{t=n}^T \varepsilon_t 1(r_3 < z_{t-j} \leq r_2) \varepsilon_{t-j} \right]^2 \\ &= T^{-2} \sum_{t,i,k,l} U_{1,t} U_{1,i} V_{1,k} V_{1,l}, \end{aligned}$$

donde $U_{1,t} = \varepsilon_t 1(r_2 < z_{t-j} < r_1) \varepsilon_{t-j}$ y $V_{1,t} = \varepsilon_t 1(r_3 < z_{t-j} < r_2) \varepsilon_{t-j}$. Ahora, si el índice mayor de t, i, k, l no es alcanzado por otro, $E(U_{1,t} U_{1,i} V_{1,k} V_{1,l}) = 0$. También, dado que los intervalos, $(r_3, r_2]$ y $(r_2, r_1]$ son disjuntos, $U_{1,i} V_{1,i} = 0$. Por lo tanto,

$$E \left[T^{-2} \sum_{t,i,k,l} U_{1,t} U_{1,i} V_{1,k} V_{1,l} \right] = T^{-2} E \left[\sum_{t,i < k} V_{1,t} V_{1,i} U_{1,k}^2 \right] + T^{-2} E \left[\sum_{t,i < k} U_{1,t} U_{1,i} V_{1,k}^2 \right].$$

Acotamos la primera esperanza, la otra se obtiene de forma similar. Así,

$$E \left[\sum_{t,i < k} V_{1,t} V_{1,i} U_{1,k}^2 \right] = E \left[\left(\sum_{t=1}^{k-1} V_{1,t} \right)^2 U_{1,k}^2 \right] = \sigma_\varepsilon^2 E \left[\left(\sum_{t=1}^{k-1} V_{1,t} \right)^2 1(r_2 < z_{k-j} < r_1) \varepsilon_{k-j}^2 \right],$$

teniendo en cuenta que $V_{1,t} V_{1,i} 1(r_2 < z_{k-j} < r_1) \varepsilon_{k-j}^2$ es una *mds* para $t > i$, excepto para $t = k - j$, es fácil probar que,

$$E \left[\left(\sum_{t=1}^{k-1} V_{1,t} \right)^2 1(r_2 < z_{k-j} < r_1) \varepsilon_{k-j}^2 \right] \leq k |r_1 - r_2| |r_2 - r_3| K_v$$

usando que $E(1(r_2 < z_t < r_1) \varepsilon_t^k / \mathfrak{S}_{t-1}) \leq |r_1 - r_2| K < \infty$, para $k = 1, 2, 3, 4$, por el supuesto A.2. Por lo tanto,

$$E \left[\sum_{t,i < k} V_{1,t} V_{1,i} U_{1,k}^2 \right] \leq k |r_1 - r_2| |r_2 - r_3| K_v \sigma_\varepsilon^2$$

Un resultado similar se puede obtener para la segunda esperanza, con lo que

$$E \left[T^{-2} \sum_{t,i,k,l} U_{1,i} U_{1,j} V_{1,k} V_{1,l} \right] \leq |r_1 - r_2| |r_2 - r_3| K_v \sigma_\varepsilon^2.$$

Con esto, se completa la prueba del lema 2.14. ■

Prueba del Teorema 2.6. La prueba de este Teorema es similar a la del Teorema 1 de Hansen (1996). Empezamos probando que $s_T(r) \Rightarrow s(r)$, donde $s(r)$ es un proceso gaussiano de media cero, y función de covarianzas, $K(r_1, r_2) = E(s(r_1)s(r_2))$. Dado que para cada r , $s_T(r)$ es una *mds* estacionaria y cuadrado-integrable, podemos aplicar el Teorema Central del Límite (TCL) puntual. Además, si suponemos que

$$T^{-1} \sum_{t=1}^T \varepsilon_t^2 h_t(\beta^0, r_1) h_t'(\beta^0, r_2) \xrightarrow{p} K(r_1, r_2) = \sigma_\varepsilon^2 E(h_t(\beta^0, r_1) h_t'(\beta^0, r_2)) = K(r_1, r_2) \quad (2.64)$$

uniformemente en r , el TCL multivariante establecerá la convergencia finito dimensional. La prueba de (2.64) se obtiene del lema 2.13, teniendo en cuenta que $\varepsilon_t^2 h_t(\beta^0, r_1) h_t'(\beta^0, r_2)$ tiene las mismas propiedades de memoria y momentos que $h_t(\beta^0, r) h_t'(\beta^0, r)$. Por lo tanto, usando el lema 2.14, que prueba la tightness del proceso $s_T(r)$, tenemos que $s_T(r) \Rightarrow s(r)$. Así, como se ha probado a lo largo de esta sección, $T^{1/2}(\hat{\beta}(r) - \beta^0) = H_T^{-1}(\beta^0, r) s_T(r) + o_p(1)$, y por lo tanto

$$T^{1/2} R'(\hat{\beta}(r) - \beta^0) \Rightarrow R' H^{-1}(\beta^0, r) s(r) = \bar{S}(r).$$

Finalmente, usando los resultados de los lemas 2.12 y 2.13 no es difícil probar que

$$\hat{\sigma}_\varepsilon^2 R' H_T^{-1}(\hat{\beta}, r) R \xrightarrow{p} \sigma_\varepsilon^2 R' H^{-1}(\beta^0, r) R$$

uniformemente en r . Por lo tanto, dado que $\hat{\alpha}(r) = R'(\hat{\beta}(r) - \beta^0) [\hat{V}(\hat{\alpha}(r))]^{-1}$ y $\hat{V}(\hat{\alpha}(r)) = \hat{\sigma}_\varepsilon^2 R' H_T^{-1}(\hat{\beta}, r) R$ para $R' = (0, 1, 0, 0)$, y usando el Teorema del Funcional Continuo,

$$G\left(\frac{T\hat{\alpha}^2(r)}{\hat{V}(\hat{\alpha}(r))}\right) \Rightarrow G\left(\frac{\bar{S}^2(r)}{\sigma_\varepsilon^2 R' H^{-1}(\beta^0, r) R}\right).$$

■

Capítulo 3

Modelos ARTIMA-Shock. El Caso No Observable.

Resumen: *En este capítulo estudiaremos el caso particular de los modelos ARTIMA-Shock, esto es, cuando la variable threshold es una función de la propia perturbación, y por lo tanto, no observable. A lo largo del capítulo nos centraremos en los casos particulares del tamaño y signo de la perturbación. Concretamente, se estudiarán las propiedades relacionadas con la persistencia e invertibilidad de la serie. Se propondrá un método de estimación de los parámetros del modelo, del que se obtendrán las condiciones suficientes para consistencia y normalidad asintótica. Igualmente, se describe la estrategia para el contraste de las hipótesis fundamentales asociadas al modelo. Finalmente, el capítulo se concluye con una aplicación de la metodología propuesta a diferentes series de índices de precios industriales.*

3.1. Introducción

En este capítulo, estudiaremos la otra clase de modelos ARTIMA tratados en la Tesis, los modelos ARTIMA-Shock. Como se comentó en el capítulo anterior, los modelos ARTIMA-Shock se caracterizan porque las asimetrías en la persistencia de las perturbaciones están generadas por alguna propiedad o característica de la propia perturbación. Esto implica que la variable threshold no es observable. El primer modelo de este tipo, con threshold en la media móvil y variable threshold no observable, fue propuesto por Wecker, y en él se permite que las perturbaciones tengan distinta persistencia en función de su signo. Wecker estudia la invertibilidad y función de autocorrelación de este tipo de series, aunque no se presenta ningún resultado de teoría asintótica respecto al método de estimación. En los posteriores trabajos con este tipo de modelos, ver por ejemplo, Elwood (1998) y Brännas y De Gooijer (1994), tampoco encontramos resultados asintóticos del método de estimación. Solo en el contexto de Inferencia Indirecta, se presentan resultados de este tipo, ver por ejemplo, Guay y Scaillet (2003) y Brännäs y Luna (1998). Sin embargo, esta metodología presenta el problema fundamental de la identificabilidad de los parámetros.

Otro modelo, también usado en el estudio de las asimetrías en la persistencia provocadas por las perturbaciones, es el modelo STOPBREAK presentado por Engle y Smith (1999). Este es un modelo de transición suave en el que es el tamaño de la perturbación el que genera las asimetrías. Engle y Smith obtienen la consistencia y normalidad asintótica del estimador máximo verosímil del modelo. Sin embargo, este modelo no es válido para estudiar aquellos casos en los que existen perturbaciones transitorias para algún tamaño de la perturbación, ya que todas las perturbaciones son permanentes. Otros modelos de transición suave en la media móvil son estudiados en Brännas, De Gooijer y Teräsvirta (1998).

Finalmente, otra aproximación diferente, cuando la variable que genera las asimetrías es no observable, es el modelo propuesto por Hamilton (1989). Sin embargo, en este tipo de modelos la variable threshold no es identificable, y por lo tanto, tampoco lo será la causa

de la asimetría, lo cual limita claramente su aplicación teórica.

Los modelos ARTIMA-Shock, que estudiaremos en este capítulo, permiten generalizar las asimetrías en la persistencia provocadas por las propias perturbaciones, permitiendo la existencia de perturbaciones transitorias y permanentes en función de alguna característica o propiedad de la propia perturbación. En principio, este capítulo se centrará en las características del signo y el tamaño, generalizando el caso del signo a valores del parámetro threshold distintos de 0. Estudiaremos las propiedades de persistencia, dando las condiciones suficientes bajo las cuales estas características pueden generar perturbaciones transitorias y permanentes. Igualmente, se estudiará la propiedad de invertibilidad y sus condiciones suficientes.

A lo largo del Capítulo se obtienen, a su vez, las condiciones suficientes para la identificabilidad de los parámetros y la validez asintótica de los contrastes propuestos. Respecto al método de estimación, como en el caso observable, estará basado en mínimos cuadrados condicionales. Sin embargo, los resultados asintóticos cambiarán radicalmente, obteniéndose la superconsistencia de todos los estimadores. Las hipótesis básicas sobre el modelo seguirán siendo la existencia de efecto threshold y la existencia perturbaciones transitorias y permanentes, para las que se propone una estrategia de contraste en dos etapas. Todos estos resultados serán aplicables al modelo propuesto por Wecker, cubriendo así la ausencia de este tipo de resultados para ese modelo. En la aplicación empírica, compararemos los resultados obtenidos por Wecker para diferentes índices de precios industriales con los obtenidos a través de la metodología propuesta en este capítulo.

El resto del capítulo se organiza como sigue. En la Sección 2 se estudian la función respuesta a un impulso y la invertibilidad de los modelos ARTIMA-Shock. En la Sección 3, se obtienen las propiedades asintóticas del método de estimación propuesto. En la Sección 4, se estudia la inferencia básica del modelo. La Sección 5, incluye la aplicación a los diferentes índices de precios industriales. La Sección 6 presenta las conclusiones. Finalmente, el Anexo

1 presenta un estudio comparativo del contraste de existencia de asimetrías propuesto. Todas las pruebas se encuentran en el Anexo 2 al final del capítulo.

3.2. Modelo ARTIMA-Shock

El modelo ARTIMA-Shock corresponde a la expresión (2.2) del Capítulo 2, donde la variable threshold es una función de la perturbación del modelo, $z_{t-1} = g(\varepsilon_{t-1})$. Como hemos comentado en la introducción, dos son las características de las perturbaciones que guiarán las asimetrías en las perturbaciones que estudiamos en este capítulo, el signo y el tamaño. Estas características se corresponderán con $g(\varepsilon_t) = \varepsilon_t$ y $r^0 = 0$, para el signo y que denominaremos modelo ARTIMA-Shock-Sign; y $g(\varepsilon_t) = |\varepsilon_t|$, para el tamaño, que denominamos ARTIMA-Shock-Size. En este capítulo, el caso del signo se generalizará al caso de $r^0 \neq 0$. Por otro lado, el caso del tamaño, $g(\varepsilon_t) = |\varepsilon_t|$, implica que la propiedad de tamaño es simétrica, esto es, una perturbación positiva es grande cuando es mayor que $r_2 > 0$, y una negativa si es menor que r_1 , con $r_1 = -r_2$. Este caso también se puede generalizar, permitiendo que $r_1 \neq -r_2$, aunque no se abordará explícitamente en este capítulo. Sin embargo, los resultados obtenidos en el mismo, son fácilmente extrapolables a dicho caso.

Como hicimos en el capítulo anterior, a continuación pasamos a enumerar los supuestos que utilizaremos a lo largo del capítulo. Así, $f(\cdot)$ será la función de densidad de ε_t , $\mathfrak{F}_{t-1} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$ el σ -algebra generado por la v.a ε hasta $t-1$, y $\theta^0 = (\theta_1^0, \theta_2^0, r^0)$ y σ_ε^2 los verdaderos valores de los parámetros del proceso.

GS.0 ε_t iid $(0, \sigma_\varepsilon)$, con una función de densidad continua $0 < f(\varepsilon_t) < \infty$ para $\forall \varepsilon_t$ y

$$E|\varepsilon_t|^{2\gamma} < \infty \text{ con } \gamma > 2.$$

GS.1 $|\theta_1^0 - \theta_2^0| = \partial^0 > 0$.

GS.2 $\Phi_p(L)$ tiene todas las raíces fuera del círculo unidad, y $\mu \in [-\bar{\mu}, \bar{\mu}]$ con $\bar{\mu} < \infty$.

I.0 $\theta_i = 1$ para algún $i = 1, 2$.

I.1 $E[\varepsilon_t | \mathfrak{S}_{t-1}] = 0$ y $E[\varepsilon_t(1 - \theta_t) | \mathfrak{S}_{t-1}] = 0$.

ASZ.0 [a] Si $|\theta_1| < |\theta_2|$, $\lambda_1^* = [\partial r f^* + E(|\theta(\varepsilon_{t-1})|)] < 1$,

[b] Si $|\theta_2| < |\theta_1|$, $\lambda_1^* = |\theta_1| + \partial r f^* < 1$

donde, $\partial = |\theta_1 - \theta_2|$, y $f^* = 2 \max_{\varepsilon} f_{\varepsilon}(\varepsilon)$

ASZ.1 [a] Si $|\theta_1^0| < |\theta_2^0|$, $\theta^0 \in \Theta_{\theta} = \{[-1 + \delta, 1 - \delta] \times [-1 + \delta, 1] \times [\underline{r}, \bar{r}]\}$ tal que

$\lambda_1(\theta) = [2\partial r f^* + \lambda_2(\theta)] < \lambda_1 < 1$ y $|\theta_1| < |\theta_2| \quad \forall \theta \in \Theta$, donde

$\lambda_2(\theta) = |\theta_1|(1 - \bar{p}(r)) + |\theta_2|\bar{p}(r)$, $\bar{p}(r) \geq \sup_k P(|\varepsilon_t + k| < r)$, y $\delta > 0$.

[b] Si $|\theta_2^0| < |\theta_1^0|$, $\theta^0 \in \Theta_{\theta} = \{[-1 + \delta, 1 - \delta]^2 \times [\underline{r}, \bar{r}]\}$ tal que

$\lambda_1(\theta) = |\theta_1| + \partial r f^* < \lambda_1 < 1$ donde $\delta > 0$ y f^* fue definida en ASZ.0.

ASZ.2 $M(\underline{r}) = E[1(|\varepsilon_{t-1}| < \underline{r})X_{t-1}X'_{t-1}] > 0$

$M(\bar{r}) = E[1(|\varepsilon_{t-1}| > \bar{r})X_{t-1}X'_{t-1}] > 0$

con $X'_{t-1} = [x_{t-1}, \dots, x_{t-p}, 1, \varepsilon_{t-1}]$ y $\underline{r} > 0$.

ASG.0 $\lambda_1^* = \max_i |\theta_i| + |\theta_1 - \theta_2| r \frac{1}{2} f^* < 1$ donde $f^* = 2 \max_{\varepsilon} f(\varepsilon)$

ASG.1 $\theta^0 \in \Theta_{\theta} = \{(-1, 1)^2 \times [\underline{r}, \bar{r}]\}$ tal que

$\lambda_1^*(\phi) = \max_i |\theta_i| + |\theta_1 - \theta_2| r f^* < \lambda_1 < 1$ donde f^* fue definida en ASG.0.

ASG.2 $M(\underline{r}) = E[1(\varepsilon_{t-1} < \underline{r})X_{t-1}X'_{t-1}] > 0$

$M(\bar{r}) = E[1(\varepsilon_{t-1} > \bar{r})X_{t-1}X'_{t-1}] > 0$

con $X'_{t-1} = [x_{t-1}, \dots, x_{t-p}, 1, \varepsilon_{t-1}]$.

Los supuestos GS.0 a GS.2 e I.0-I.1 coinciden en motivación y contenido con los utilizados en el modelo ARTIMA-Observable. Los supuestos ASZ.0 y ASG.0, establecen las

condiciones suficientes para la invertibilidad del modelo en el caso del signo y el tamaño respectivamente. Los supuestos ASZ.1 y ASG.1 establecen las condiciones suficientes para la consistencia e inferencia asintótica en el caso del signo y el tamaño. Finalmente, las condiciones ASZ.2 y ASG.2 representan las condiciones de rango completo para ambos casos y garantizan la identificabilidad de los parámetros. Como es fácilmente comprobable, bajo las condiciones GS.0 a GS.2 se cumplirá la condición A.2 utilizada en el modelo ARTIMA-Observable, y que no tiene correspondencia en los supuestos aquí presentados.

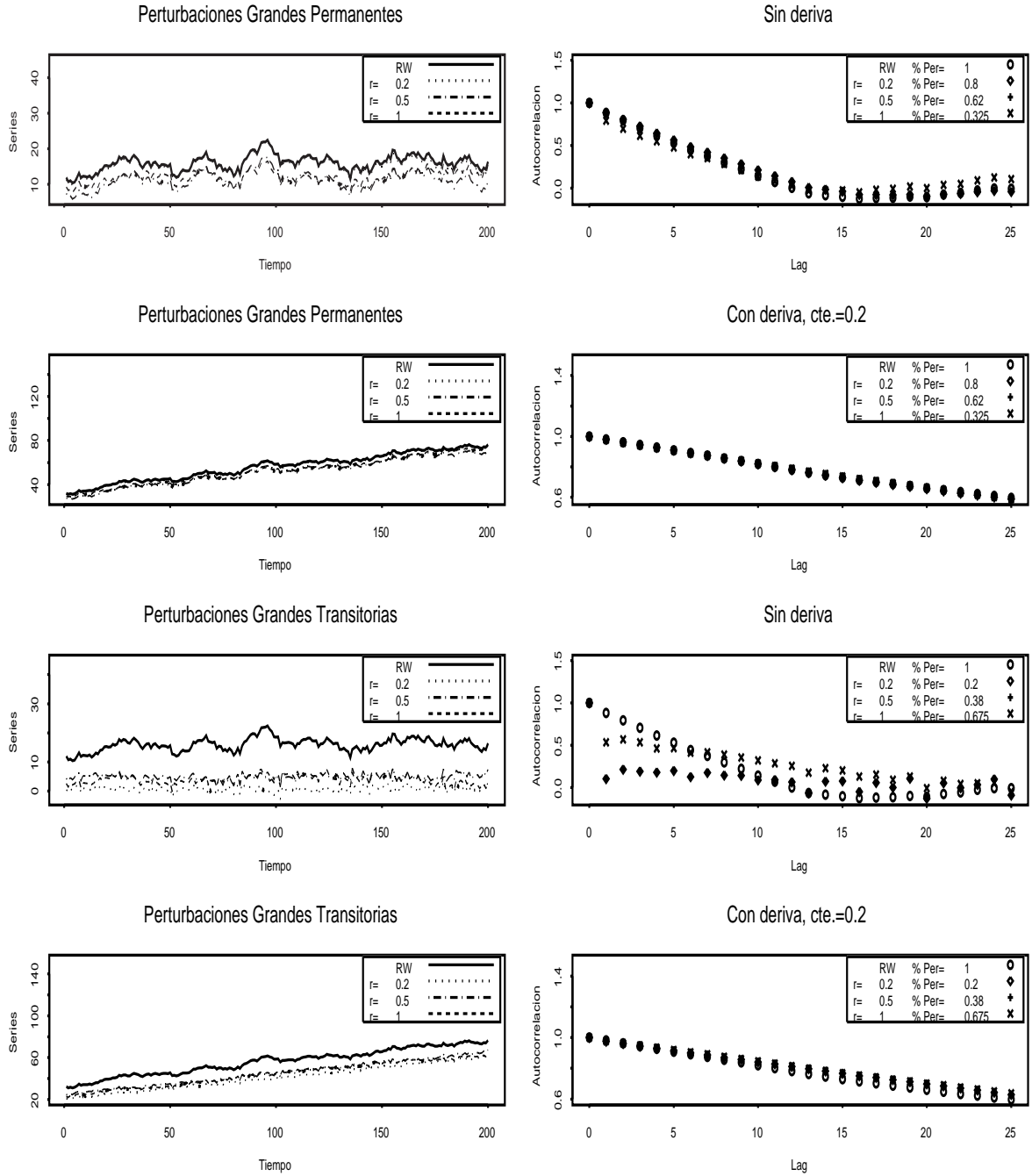
En las figuras 3.1 y 3.2, se encuentran representadas varias series correspondientes a modelos ARTIMA-Shock, incluyendo en cada caso un paseo aleatorio. En los gráficos del lado izquierdo se muestran las series para distintos valores de r , θ_i y μ , y en el derecho sus funciones de autocorrelación. Concretamente, los modelos son:

Figura 2.1: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r)$

Figura 2.2: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(\varepsilon_{t-1} > r) - \theta_2 \varepsilon_{t-1} 1(\varepsilon_{t-1} < r)$

Los valores que pueden tomar θ_1 y θ_2 son 0 y 1. Si $\theta_i = 1$ denotaremos al régimen i como transitorio, y si $\theta_i = 0$ como permanente. Como en el Capítulo 2, en el gráfico que representa la autocorrelación se muestra el porcentaje de perturbaciones en cada uno de los regímenes, y como sucedía entonces, la función de autocorrelación del paseo aleatorio no muestra siempre el decaimiento más lento, a pesar de ser la serie con mayor porcentaje de perturbaciones permanentes. De nuevo, se observa que en el caso en que las perturbaciones permanentes son las de mayor varianza, la tasa de decaimiento es similar. Esta diferencia en el decaimiento, desaparece cuando se introduce la deriva.

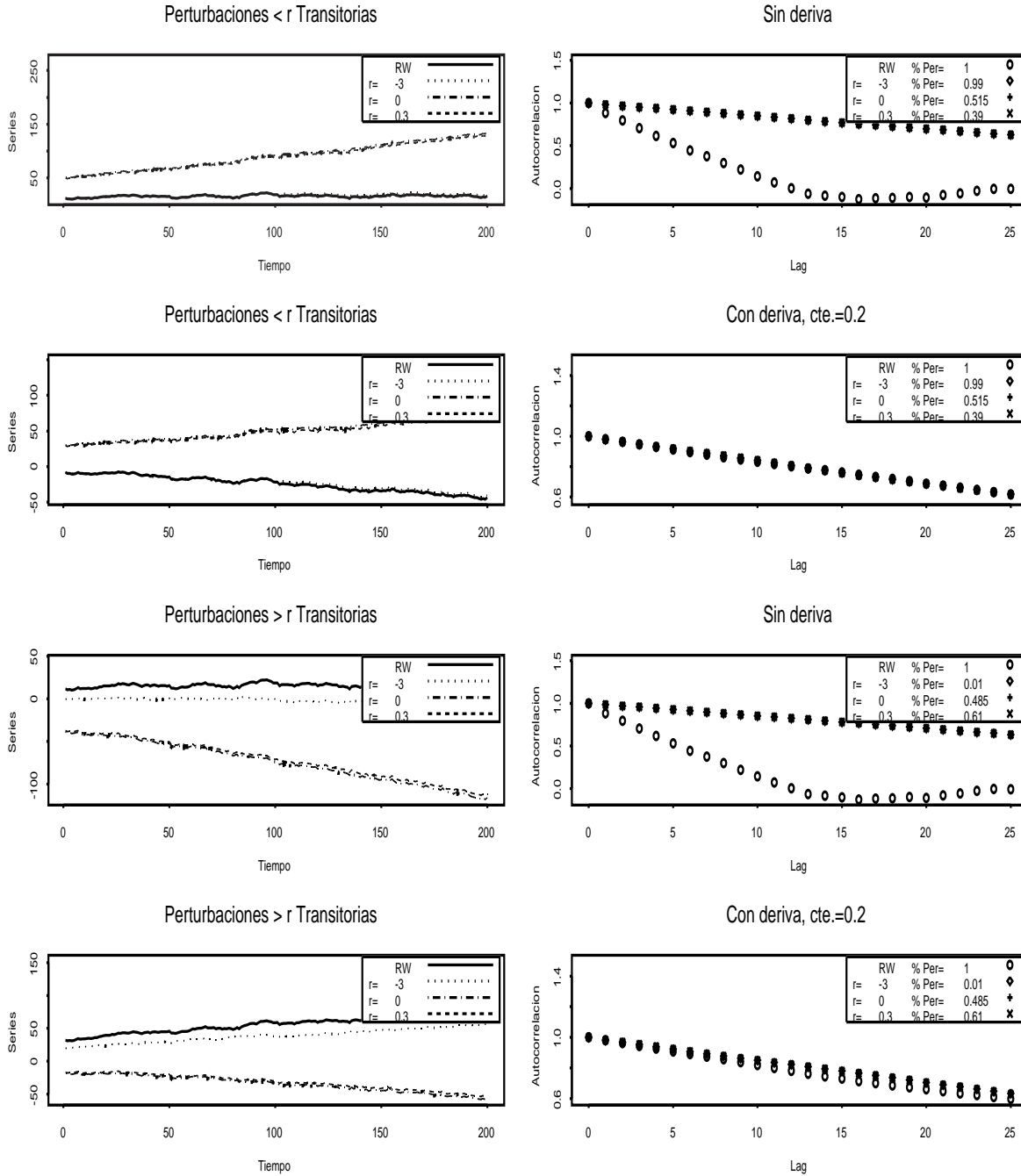
Después de ver los supuestos suficientes para los resultados que se obtendrán en adelante y algunos gráficos de las series, pasamos a estudiar la función respuesta a un impulso generalizada e invertibilidad de los modelos ARTIMA-Shock.

Figura 3.1: Serie ARTIMA con $z_t = |\varepsilon_t|$ 

Nota: ARTIMA: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r)$; RW: $(1 - L)y_t = \mu + \varepsilon_t$.

Régimen Transitorio: $\theta_i = 1$; Régimen Permanente: $\theta_i = 0$.

Figura 3.2: Serie ARTIMA con $z_t = \varepsilon_t$



Nota: ARTIMA: $(1 - L)y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(\varepsilon_{t-1} > r) - \theta_2 \varepsilon_{t-1} 1(\varepsilon_{t-1} < r)$; RW: $(1 - L)y_t = \mu + \varepsilon_t$.

Régimen Transitorio: $\theta_i = 1$; Régimen Permanente: $\theta_i = 0$.

3.2.1. Función Respuesta a un Impulso.

En el capítulo introductorio se discutieron la propiedades de la GI para los modelos no lineales, y en el capítulo anterior, se demostró en el lema 2.1 que bajo los supuestos I.0 e I.1, su expresión para los modelos ARTIMA viene dada por:

$$GI(k, v_t, \mathfrak{S}_{t-1}) = \begin{cases} \psi_{k-1}\varepsilon_t & \text{si } \theta_t = 1 \\ \left[(1 - \theta) \sum_{j=0}^{k-2} \psi_j + \psi_{k-1} \right] \varepsilon_t & \text{si } \theta_t = \theta \neq 1. \end{cases}$$

Para el caso de los modelos ARTIMA-Shock, $v_t = \varepsilon_t$ y $\mathfrak{S}_{t-1} = \sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots)$. Como es fácilmente comprobable, para los modelos ARTIMA-Shock-Size, una condición suficiente para I.1 es que ε_t sea *iid* con función de densidad simétrica respecto a 0. Sin embargo, para el caso de ARTIMA-Shock-Sign nunca se cumplirá la condición I.1 para $\theta_1 \neq \theta_2$, y por lo tanto, no será posible obtener, en función de la GI, perturbaciones transitorias. Para el caso del signo se obtiene que $E[\varepsilon_t(1 - \theta_t)/\mathfrak{S}_{t-1}] = \mu_E \neq 0$, con lo que la perturbación en el régimen no invertible no es puramente transitoria, aunque no forma parte del componente permanente en la descomposición de Beveridge y Nelson de la serie, como ya se discutió en el capítulo introductorio.

Como en el caso Observable, conviene destacar que la selección de ε_{t-1} como variable threshold garantiza que sean el signo o tamaño de la propia perturbación, ε_t , quien determine su persistencia.

Por último, y como hicimos en la correspondiente sección del capítulo anterior, presentamos a continuación la estimación de la función GI para una serie de modelos ARTIMA-Shock, TAR y el paseo aleatorio. El método de estimación es el mismo que entonces, y los modelos seleccionados, a parte del paseo aleatorio, son:

$$\mathbf{ARTIMA\ 1} : (1 - 0.5L)(1 - L)y_t = \varepsilon_t - 0.5\varepsilon_{t-1}1(|\varepsilon_{t-1}| > 0.6) - \varepsilon_{t-1}1(|\varepsilon_{t-1}| < 0.6).$$

$$\mathbf{TAR\ 1} : y_t = y_{t-1}1(|\varepsilon_{t-1}| > 0.6) + 0.5y_{t-1}1(|\varepsilon_{t-1}| < 0.6) + \varepsilon_t.$$

ARTIMA 2 : $(1 - 0.5L)(1 - L)y_t = \varepsilon_t - 0.5\varepsilon_{t-1}1(\varepsilon_{t-1} > 0) - \varepsilon_{t-1}1(\varepsilon_{t-1} < 0)$.

TAR 2 : $y_t = y_{t-1}1(\varepsilon_{t-1} > 0) + 0.5y_{t-1}1(\varepsilon_{t-1} < 0) + \varepsilon_t$.

En todos los procesos, ε_t es *iid* $N(0, 1)$. En el primer grupo de modelos, ARTIMA 1 y TAR 1, la asimetría en la persistencia viene provocada por el tamaño de la perturbación, mientras que en el segundo grupo, por el signo. Como es fácilmente comprobable, dado que ε_t es *iid*, para los procesos TAR1 y TAR2, se tiene que:

$$GI(k, \varepsilon_t, w_{t-1}) = E(\phi_t / \varepsilon_t) E(\phi_t)^{k-1} \varepsilon_t,$$

con $\phi_t = \theta_1 1(g(\varepsilon_t) > r) + \theta_2 1(g(\varepsilon_t) < r)$. Las figuras 3.3 y 3.4 muestran los resultados para el caso del tamaño. En la figura 3.3 se presentan los resultados de estimación de la GI a través de la función de densidad condicionada al tamaño, para cada uno de los tres modelos. Para el modelo ARTIMA, y como se obtiene del lema 2.1, la dispersión de la densidad de la GI condicionada a las perturbaciones pequeñas converge a 0, dado que $\theta_2 = 1$. Para las perturbaciones grandes, sin embargo, la dispersión se mantiene constante. Para el caso TAR, y también de acuerdo al resultado teórico, la dispersión de la GI converge a 0 en ambos casos. Por el contrario, en el caso del paseo aleatorio, la dispersión de la GI se mantiene constante. Además, y como se muestra en el gráfico 3.4, en este caso no hay asimetrías de ningún tipo en la persistencia, consecuencia obvia de la linealidad del modelo. Sin embargo, para el caso TAR, si hay asimetrías en términos de la tasa de convergencia media. Como en el Capítulo anterior, la figura 3.4 representa la mediana de las GI condicionadas y estandarizadas por ε_t , y en ella podemos ver como la respuesta media de una perturbación grande en el modelo ARTIMA es igual a la obtenida en el paseo aleatorio, mientras que la de las perturbaciones pequeñas se asemeja a la de un proceso autoregresivo estacionario.

Las figuras 3.5 y 3.6 muestran los resultados para el caso del signo. En este caso, el régimen de MA no invertible corresponde a las perturbaciones negativas, lo que implica,

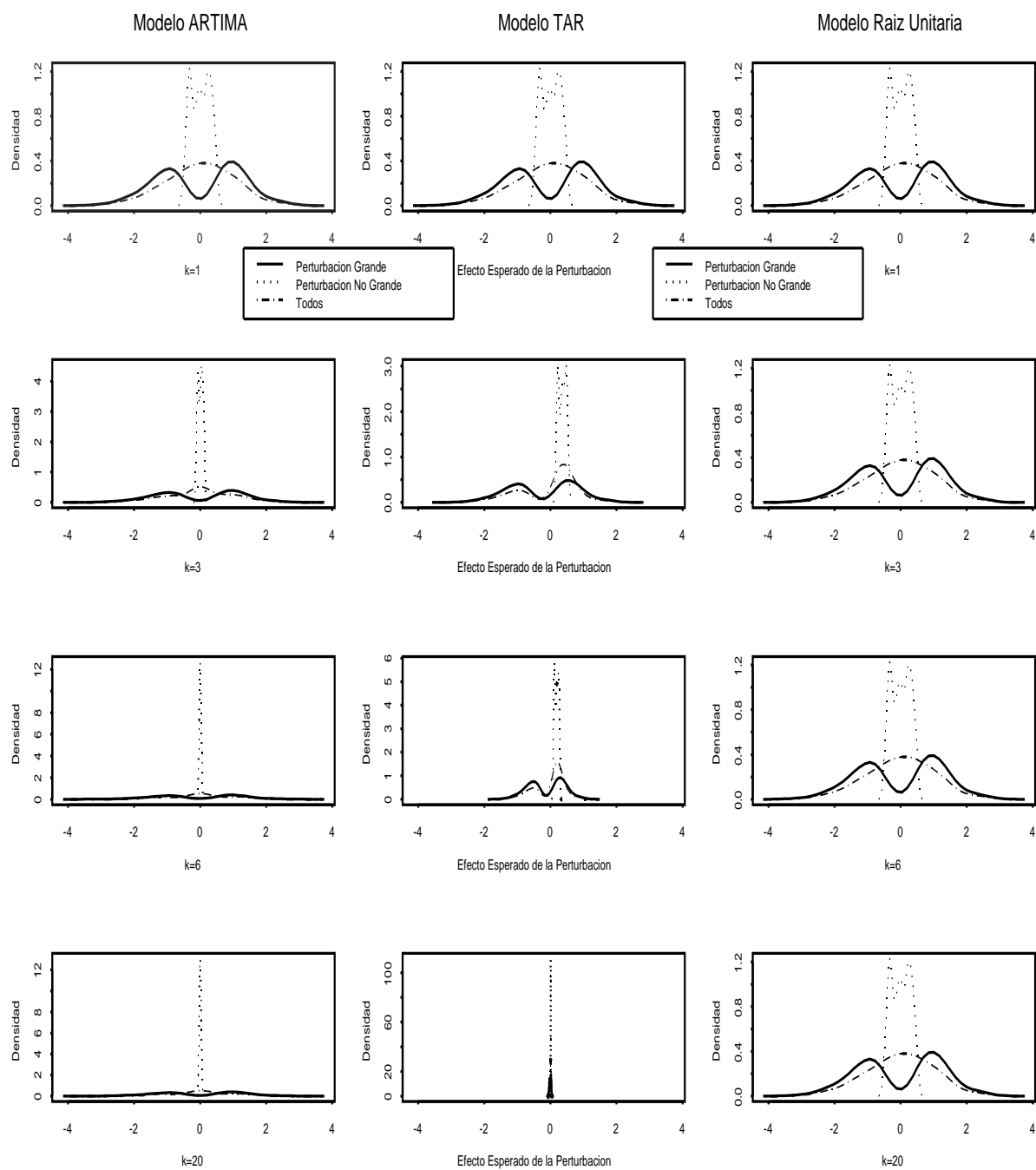
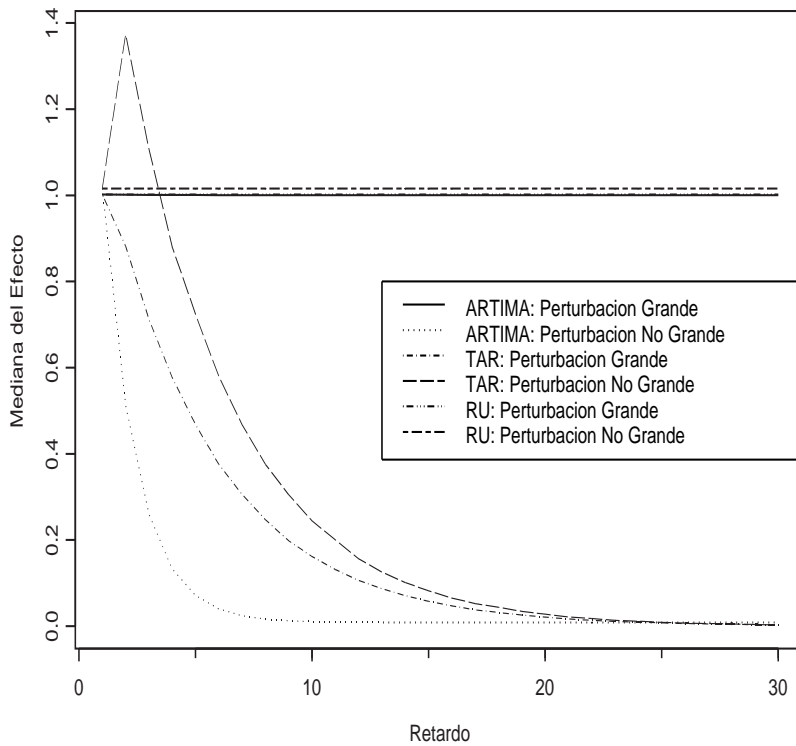
Figura 3.3: Densidad Estimada de la GI. Caso $|\varepsilon_t|$.

Figura 3.4: Esperanza Estimada de la $\frac{GI}{\varepsilon_t}$. Caso $|\varepsilon_t|$.



como puede verse en la figura 3.5, que la dispersión de la GI condicionada a este tipo de perturbación converge a 0. Sin embargo, como refleja la figura 3.6 la densidad no esta centrada en 0. El resto de resultados para los otros dos modelos son similares a los obtenidos en el caso anterior.

3.2.2. Invertibilidad

En esta sección discutiremos las condiciones bajo las cuales los procesos ARTIMA-Shock son invertibles, según la definición de invertibilidad para modelos no lineales introducida por Granger y Andersen, y recogida en el capítulo anterior en la definición 2.1. Como vimos en la Sección 2.3.2, esta definición garantiza la recuperación de las perturbaciones a través

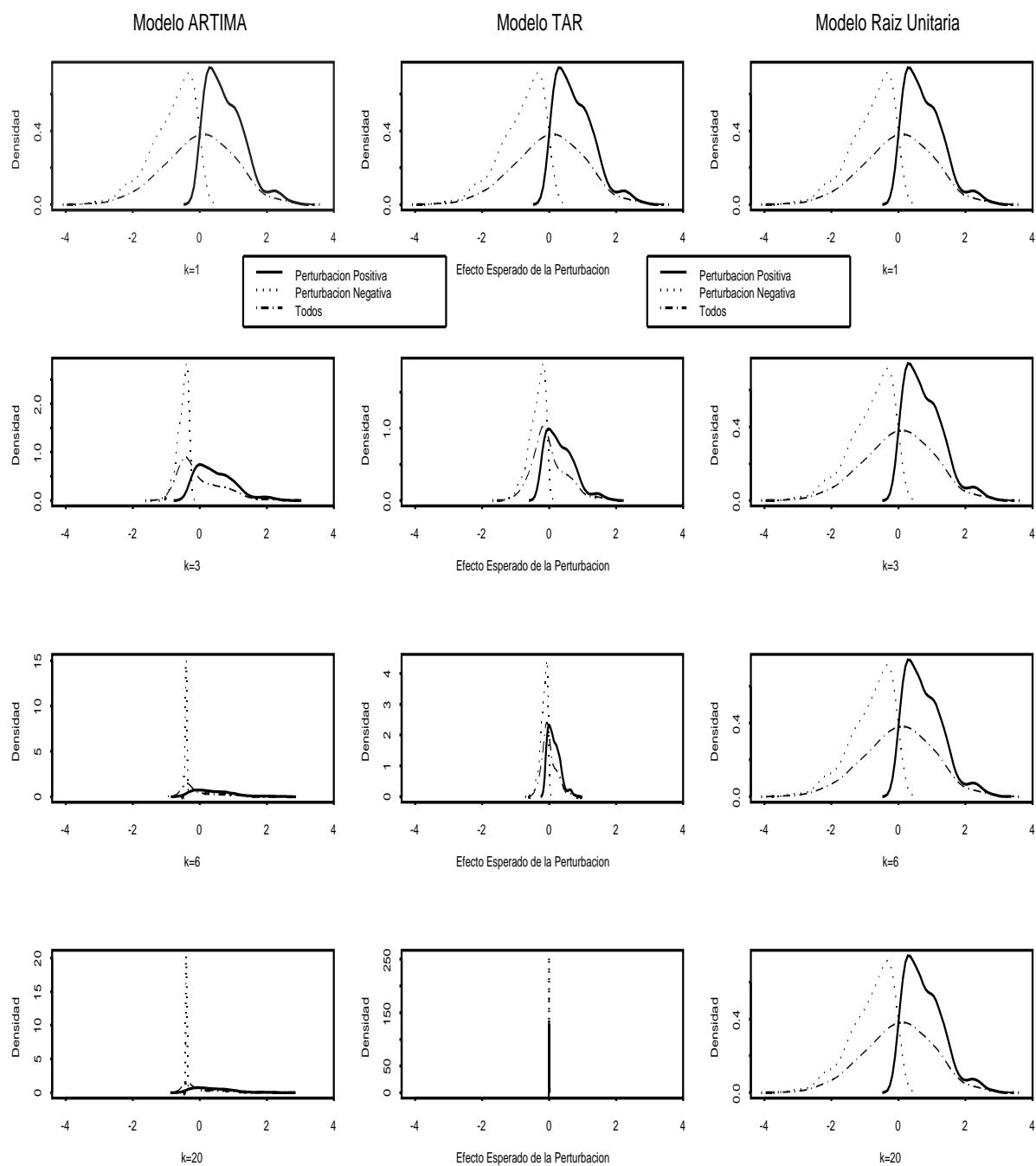
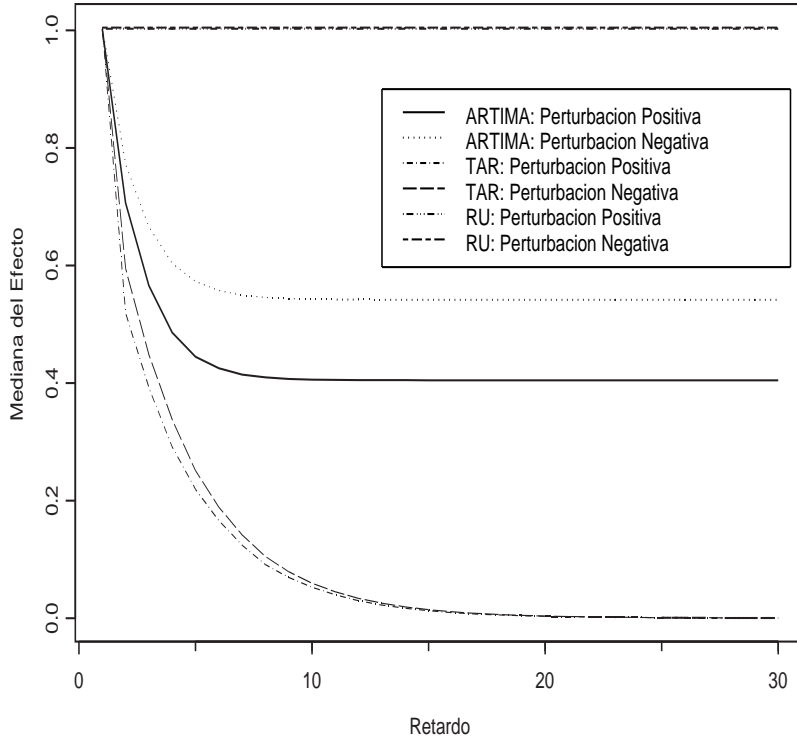
Figura 3.5: Densidad Estimada de la GI. Caso ε_t .

Figura 3.6: Esperanza Estimada de la $\frac{GI}{\varepsilon_t}$. Caso ε_t .



de la serie observada, y da sentido, por tanto, a la predicción en el contexto de MA y a la definición de \mathfrak{S}_{t-1} en la GI en función de ε_t .

Los principales resultados sobre invertibilidad para los modelos ARTIMA-Shock se recogen en los siguientes lemas.

Lema 3.1 *Sea y_t un proceso ARTIMA-Shock-Size; entonces bajo los supuestos G.0, G.2 y ASZ.0 a) o b), y_t será invertible.*

Prueba, ver Anexo 2.

Lema 3.2 *Sea y_t un proceso ARTIMA-Shock-Sign; entonces bajo los supuestos G.0, G.2 y ASG.0, y_t será invertible.*

Prueba, ver Anexo 2.

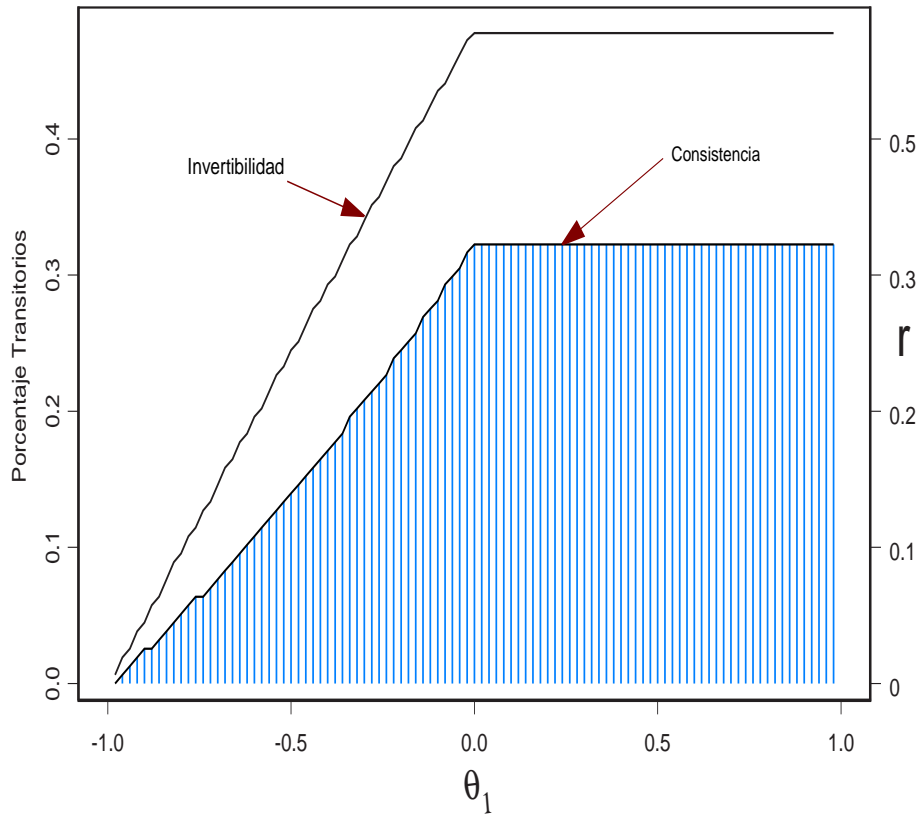
Estudiando en detalle los supuestos ASZ.0 Y ASG.0, podemos distinguir dos partes. La primera hace referencia a $E(|\theta(\varepsilon_{t-1})|)$ o $\max_i \theta_i$, que debe ser menor que 1. Esta condición elimina la posibilidad de sobrediferenciación, fundamentalmente cuando permitimos la posibilidad de una raíz unitaria en alguno de los regímenes MA. Sin embargo, la no invertibilidad no es solo una cuestión de sobrediferenciabilidad, si no que también es una cuestión de no linealidad y discontinuidad. La segunda parte del supuesto, $\partial r f^*$, controla el grado de discontinuidad, medido como el producto del tamaño de la discontinuidad, ∂r , y la probabilidad de que esta se produzca, f^* . Así, por ejemplo, en el caso del signo, si $r = 0$, el proceso será continuo en todos los parámetros, y la condición se reduce al supuesto de sobrediferenciación, que coincide con la condición establecida por Wecker (1981). Para el caso del tamaño, no existe la posibilidad de continuidad, de forma que la discontinuidad del modelo ha de ser acotada si se quiere preservar la condición de invertibilidad.

La condición ASZ.0 en el caso del tamaño, sólo permite la raíz unitaria en el régimen de MA asociado a las perturbaciones pequeñas. Luego, sólo el caso de perturbaciones pequeñas transitorias admite una representación invertible. En el caso del signo, la invertibilidad del proceso no permite la existencia de raíces unitarias en ninguno de los regímenes. Sin embargo, como ya comentamos en el apartado anterior, la existencia de esta raíz unitaria en algún régimen MA, no garantizaba la existencia de perturbaciones transitorias puras.

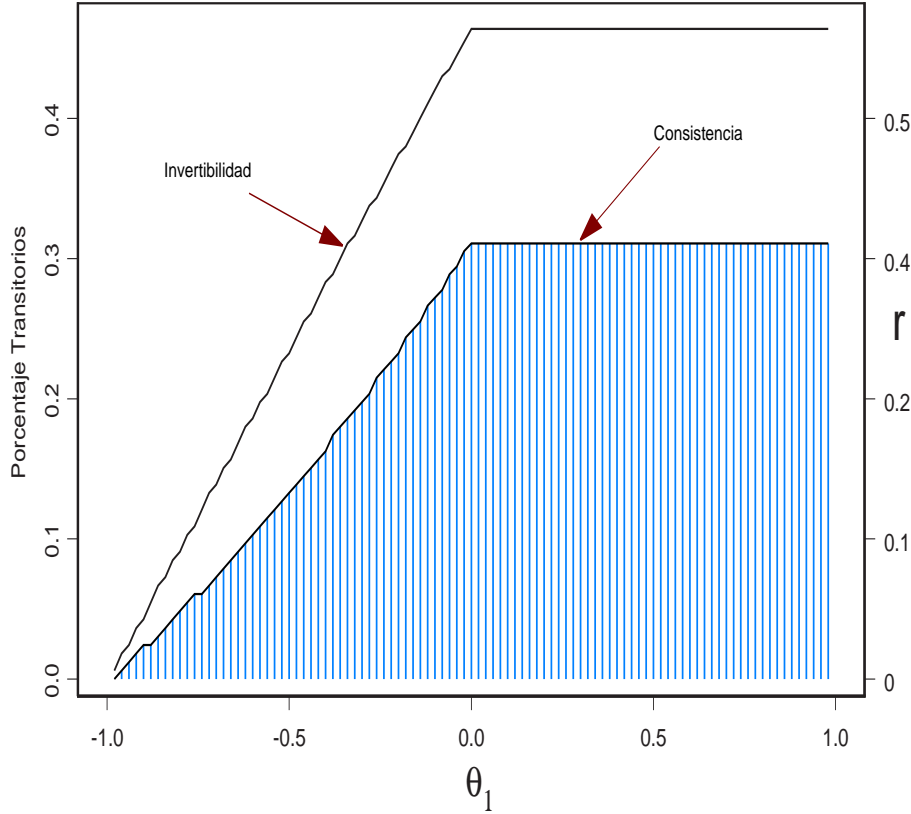
Respecto a las condiciones ASZ.0 y ASG.0, conviene comentar que en el caso del signo, con $r \neq 0$, la restricción de no linealidad aparece en función de $1/2f^*$, ya que en el modelo hay dos regímenes en la MA. En el caso del tamaño, la restricción de no linealidad aparece en función de f^* , ya que son tres los regímenes en la MA. Así, en general, cuando tengamos $m + 1$ regímenes, la restricción de no linealidad aparecerá en función de $m/2f^*$.

Finalmente, en las Figuras 3.7 y 3.8 se representan los valores de θ_1 , r y el porcentaje de perturbaciones transitorias que satisfacen las condiciones ASZ.0 y ASZ.1 cuando $\theta_2 = 1$

Figura 3.7: Espacio Paramétrico Invertible y Consistente con $\varepsilon_t \sim N(0, 1)$



y la función de densidad de ε_t es una $N(0, 1)$, Figura 3.7, y una t -Student de 5 grados de libertad, figura 3.8. La línea superior, marca la frontera para la invertibilidad, el área rallada, el espacio paramétrico para la consistencia. En el caso de la $N(0, 1)$, la invertibilidad permitirá más de un 50 % de perturbaciones transitorias, y la consistencia más de un 30 %. En el caso de una t -Student, los porcentajes de perturbaciones transitorias disminuye, aunque el valor de r máximo admitido aumenta. Se puede comprobar, que en ambos casos, el porcentaje de perturbaciones transitorias es independiente de σ_ε^2 .

Figura 3.8: Espacio Paramétrico Invertible y Consistente con $\varepsilon_t \sim t_5$ 

3.3. Estimación

En esta sección se presentan los resultados de identificabilidad, consistencia y tasa de convergencia del método de estimación propuesto para los modelos ARTIMA-Shock. Como en el caso Observable, las pruebas se obtendrán para el modelo,

$$(1 - \rho L) \Delta y_t = (1 - \rho L) x_t = \mu + \varepsilon_t - \theta_{t-1} \varepsilon_{t-1} = \mu + \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & \text{si } g(\varepsilon_{t-1}) > r \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & \text{si } g(\varepsilon_{t-1}) < r \end{cases} \quad (3.1)$$

para $g(\varepsilon_{t-1}) = \varepsilon_{t-1}$ y $g(\varepsilon_{t-1}) = |\varepsilon_{t-1}|$. Las pruebas para el caso más general, $p > 1$ se obtienen fácilmente de las aquí presentadas.

En este caso, la función objetivo se puede expresar como,

$$Q_T(\beta, r) = \sum_{t=1}^T e_t^2(\beta, r)$$

con $\beta = (\theta_1, \theta_2, \rho, \mu)$ y

$$e_t(\beta, r) = x_t - \rho x_{t-1} - \mu + \theta_{t-1} e_{t-1}(\beta, r) \quad e_0 = 0$$

$$\theta_{t-1} = \begin{cases} \theta_1 & \text{if } g(e_{t-1}) > r \\ \theta_2 & \text{if } g(e_{t-1}) < r \end{cases}.$$

Definimos, por tanto, el estimador de mínimos cuadrados condicionales (CLSE) de ϕ^0 , con $\phi^0 = (\beta^0, r^0)$ el verdadero valor de los parámetros, como $\hat{\phi}_T = (\hat{\beta}, \hat{r})$, que debe satisfacer:

$$\hat{\beta}_T(r) = \arg \min_{\beta \in \Theta_\beta} Q_T(\beta, r)$$

y

$$\hat{r}_T = \arg \min_{r \in [\underline{r}, \bar{r}]} Q_T(\hat{\beta}_T(r)),$$

donde $\Theta_\beta = \{\phi \quad t.q \quad \theta \in \Theta_\theta, \rho \in (-1, 1), \mu \in [-\bar{\mu}, \bar{\mu}]\}$ y Θ_θ está definido en los supuestos ASG.1 y ASZ1 para el caso del signo y del tamaño respectivamente. La diferencia fundamental con el caso observable, es que la función objetivo no será derivable con respecto a los errores e_t , ya que estos entran a través de la función indicadora, $1(g(e_t) > r)$. Por lo tanto, la función objetivo no será derivable respecto a ningún parámetro. Esto marcará las diferencias en los resultados sobre la tasa de convergencia de los estimadores. Estos resultados se muestran en los siguientes dos Teoremas. Así, para el caso del tamaño, tenemos el siguiente resultado.

Teorema 3.1 *Bajo las condiciones G.0-G.3, y ASZ.1-ASZ.2, $\hat{\phi}_{i,T} = \phi_i^0 + O_p(T^{-1})$ para $i = 1, \dots, 5$.*

Prueba, ver Anexo 2.

Este Teorema, establece las condiciones suficientes para la identificabilidad y consistencia de los estimadores en el caso de los modelos ARTIMA-Shock-Size. Si comparamos este resultado con el obtenido en el Teorema 2.2, vemos que los resultados son muy diferentes, ya que la tasa de convergencia de estos estimadores es en todos los casos T . Por lo tanto, todos los estimadores son superconsistentes. Como comentamos en el párrafo anterior, la razón de este resultado se encuentra en que, en este caso, todos los parámetros comparten la no linealidad del parámetro r a través de la función indicador. Como veremos en la sección siguiente, este resultado marcará la estrategia a seguir para realizar la inferencia sobre los parámetros del modelo.

Para el caso de los modelos ARTIMA-Shock-Sign, los resultados se muestran en el siguiente Teorema.

Teorema 3.2 *Bajo las condiciones $G,0-G,3$, y $ASG,1$,*

a) *Si $r^0 = 0$ y conocido, $\hat{\phi}_{i,T} = \phi_i^0 + O_p(T^{-1/2})$ para $i = 1, \dots, 4$.*

b) *Si $|r^0| > 0$, $\hat{\phi}_{i,T} = \phi_i^0 + O_p(T^{-1})$ para $i = 1, \dots, 5$.*

Prueba, ver Anexo 2.

De nuevo, este Teorema establece la identificabilidad de los parámetros y la consistencia de todos los estimadores. Respecto a la tasa de convergencia de los mismos, se distinguen dos casos. Como vimos en la sección de invertibilidad, en el caso del signo, cuando $r^0 = 0$ y conocido, el proceso es continuo en los parámetros, aunque no diferenciable. Por lo tanto, en este caso, la tasa de convergencia coincide con el caso Observable, que también presenta continuidad en todos los parámetros excepto r^0 . Sin embargo, cuando $|r^0| > 0$ el proceso será discontinuo, y lo será respecto a todos los parámetros. En este caso, el resultado es el mismo que en el caso del tamaño.

3.4. Inferencia

Como en el caso Observable, las hipótesis básicas de los modelos ARTIMA-Shock son dos, la primera, la existencia de efecto threshold, y la segunda, la existencia de asimetrías en términos de perturbaciones transitorias y permanentes. Por lo tanto, en una primera etapa contrastamos la hipótesis nula de no existencia de threshold. Cuando rechacemos esta hipótesis nula, y sea factible un régimen MA no invertible en el modelo, pasaremos a contrastar la hipótesis nula de existencia de perturbaciones transitorias en dicho régimen.

En lo que resta de sección, dividiremos los resultados de inferencia en dos casos, según r^0 sea conocido o no.

3.4.1. Inferencia para r^0 conocido

Como se comento en el capítulo anterior para el caso Observable, este no es el escenario más probable, exceptuando el caso del signo con $r^0 = 0$. Además, para los modelos ARTIMA-Shock con $|r^0| > 0$, conocer el valor de r^0 no modifica la teoría asintótica (al contrario de lo que ocurre en el caso Observable), ya que la función objetivo a minimizar sigue sin ser continua en el resto de los parámetros. Por esta razón, en esta sección sólo abordaremos el caso de ARTIMA-Shock-Sign con $r^0 = 0$ conocido.

El siguiente Teorema establece los supuestos para la normalidad asintótica de los estimadores. Este resultado será suficiente para la realización de los dos contrastes anteriormente descritos, al igual que sucedía en el caso observable.

Teorema 3.3 *Sea y_t un proceso ARTIMA-Shock-Sign con $r^0 = 0$ conocido, entonces bajo G.0, G.2, G.3, ASG.1 y ASG.2,*

$$T^{1/2} \left(\hat{\beta} - \beta \right) \xrightarrow{d} N \left(\vec{0}, \sigma_\varepsilon^2 H^{-1} \right)$$

con,

$$\begin{aligned}
 H &= \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T E [h_t(\beta^0) h_t'(\beta^0)] \\
 h_t &= \left(h_t^{(1)}, h_t^{(2)}, h_t^{(3)}, h_t^{(4)} \right)' \\
 h_t^{(1)} &= \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^+, \quad h_t^{(2)} = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^-, \\
 h_t^{(3)} &= - \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 x_{t-1-j}, \quad h_t^{(4)} = - \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \\
 \theta_t^0 &= \theta_1^0 1(\varepsilon_t > 0) + \theta_2^0 1(\varepsilon_t < 0) \\
 \varepsilon_t^+ &= 1(\varepsilon_t > 0) \varepsilon_t \quad y \quad \varepsilon_t^- = \varepsilon_t - \varepsilon_t^+.
 \end{aligned}$$

Prueba, ver Anexo 2.

Como se puede comprobar, la distribución de los estimadores es asintóticamente normal de media 0, con matriz de varianzas y covarianzas similar a la obtenida en el caso Observable, diferenciándose en la definición de ε_t^+ , ε_t^- y θ_t^0 . Como en aquel caso, para hacer factible la inferencia, es necesario obtener una estimación consistente de dicha matriz de varianzas. En el siguiente Teorema se establece el estimador de la matriz de varianzas y covarianzas y su consistencia bajo las condiciones del Teorema 3.3.

Teorema 3.4 *Bajo las condiciones del Teorema 3.3*

$$\hat{\sigma}_\varepsilon^2 \frac{1}{T} \sum_{t=2}^T \hat{H}_t^{i,j} \xrightarrow{p} \sigma_\varepsilon^2 H^{i,j}$$

con

$$\begin{aligned}
 \hat{\sigma}_\varepsilon^2 &= \frac{1}{T} \sum_{t=1}^T \hat{\varepsilon}_t^2 \\
 \hat{H}_t^{i,j} &= \hat{h}_t^{(i)} \hat{h}_t^{(j)} \\
 \hat{h}_t^{(1)} &= \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \hat{\varepsilon}_{t-1-j}^+, \quad \hat{h}_t^{(2)} = \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \hat{\varepsilon}_{t-1-j}^-, \\
 \hat{h}_t^{(3)} &= - \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 x_{t-1-j}, \quad \hat{h}_t^{(4)} = - \sum_{j=0}^t \prod_{k=1}^j \hat{\theta}_{t-k}^0 \\
 \hat{\varepsilon}_t &= e_t(\hat{\beta}); \quad \hat{\varepsilon}_t^+ = \hat{\varepsilon}_t 1(\hat{\varepsilon}_t > 0); \quad \hat{\varepsilon}_t^- = \hat{\varepsilon}_t - \hat{\varepsilon}_t^+ \\
 \hat{\theta}_t &= \hat{\theta}_1 1(\hat{\varepsilon}_t > 0) + \hat{\theta}_2 1(\hat{\varepsilon}_t < 0).
 \end{aligned}$$

y $\hat{\beta}$ el estimador CLS.

Prueba, ver Anexo 2.

Apoyados en los resultados de estos dos Teoremas, para realizar el contraste de existencia de efecto threshold proponemos el siguiente estadístico:

$$W_{0,T} = \frac{T \left(R' \hat{\beta} \right)^2}{\left[\hat{\sigma}_\varepsilon^2 \left(R' \hat{H}^{-1} R \right) \right]},$$

con $R' = (1, -1, 0, 0)$ y que bajo $H_0 : \theta_1 = \theta_2$ convergerá a una χ_1^2 . Este estadístico coincide con el propuesto por Brännas y De Gooijer (1994), aunque en su trabajo no se justifican los resultados asintóticos.

En el caso del signo, la condición de invertibilidad excluye la posibilidad de $\theta_i = 1$ para algún régimen de la MA, con lo que, en principio, el contraste de existencia de perturbaciones transitorias basado en la hipótesis nula $\theta_i = 1$ no será válido asintóticamente.

3.4.2. Inferencia cuando r^0 es desconocido

En esta sección presentamos los resultados asintóticos que nos permitirán realizar los contrastes de hipótesis cuando r^0 es desconocido. Al igual que en el caso Observable, em-

pezaremos asumiendo que existen asimetrías en la persistencia y calculando la distribución asintótica de los estimadores.

Como se ha visto a lo largo del capítulo, si $|r^0| > 0$ la función objetivo no es continua en ninguno de los parámetros, lo cual dificulta la obtención de la distribución asintótica. Por lo tanto, aprovechando la superconsistencia de los estimadores, se propone estimar los parámetros en dos etapas. En el siguiente Teorema se prueba que cuando r^0 es desconocido, pero $|r^0| > 0$, la distribución asintótica de los estimadores de la segunda etapa para $\beta^0 = (\theta_1^0, \theta_2^0, \rho^0, \mu^0)$, es normal multivariante de media cero.

El procedimiento será el siguiente. En la primera etapa, estimamos los parámetros por el método CLS descrito en la sección de estimación. De esta primera estimación, obtenemos,

$$\hat{\varepsilon}_t = x_t - \hat{\rho}x_{t-1} - \hat{\mu} + \hat{\theta}_{t-1} (g(\hat{\varepsilon}_{t-1})) \hat{\varepsilon}_{t-1},$$

que usaremos como regresor en la estimación de la segunda etapa, con $g(\hat{\varepsilon}_t) = |\hat{\varepsilon}_t|$ en el caso del tamaño, y $g(\hat{\varepsilon}_t) = \hat{\varepsilon}_t$ en el caso del signo. En esta segunda etapa, se propone la estimación por Mínimos Cuadrados Ordinarios (MCO) de la siguiente ecuación

$$x_t = \alpha_1 + \alpha_2 x_{t-1} + \alpha_3 \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) > \hat{r}) + \alpha_4 \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) < \hat{r}) + u_t. \quad (3.2)$$

El siguiente Teorema establece la normalidad asintótica de los estimadores MCO de esta ecuación.

Teorema 3.5 *Si $|r^0| > 0$ y bajo los supuestos del Teorema 3.1 y 3.2 para el caso del tamaño y del signo respectivamente,*

$$T^{-1/2} (\hat{\alpha} - \alpha^0) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 E \left(Z_t^{0'} Z_t^0 \right)^{-1} \right),$$

con $\hat{\alpha}$ el estimador MCO de la ecuación (3.2), $Z_T^{0'} = (Z_{T,1}^0, \dots, Z_{T,T-1}^0)$,

$Z_{T,t-1}^{0'} = (1, x_{t-1}, \varepsilon_{t-1} 1(g(\varepsilon_{t-1}) > r^0), \varepsilon_{t-1} 1(g(\varepsilon_{t-1}) < r^0))$, y $\alpha^{0'} = (\mu^0, \rho^0, \theta_1^0, \theta_2^0)$.

Prueba, ver Anexo 2.

Para el caso del tamaño, como quedó probado en el Lema 2.1, las perturbaciones pequeñas pueden ser transitorias y el modelo seguir siendo invertible. Por lo tanto, toda la teoría asintótica desarrollada en esta sección es válida para contrastar la hipótesis de perturbaciones pequeñas transitorias. Para ello planteamos el siguiente contraste:

$$H_0: \theta_2 = 1 \Leftrightarrow \alpha_4 = 1$$

$$H_a: \theta_2 \neq 1 \Leftrightarrow \alpha_4 \neq 1.$$

El estadístico propuesto para este contraste es,

$$W_{1,T} = \frac{T(\hat{\alpha}_4 - 1)^2}{\left[\hat{V}(\hat{\beta})[4, 4]\right]},$$

con

$$\begin{aligned} \hat{V}(\hat{\beta}) &= \hat{\sigma}_\varepsilon^2 T^{-1} \left(\hat{Z}'_T \hat{Z}_T \right)^{-1} \\ \hat{Z}'_{T,t} &= (1, x_{t-1}, \hat{\varepsilon}_{t-1} 1(|\hat{\varepsilon}_{t-1}| > \hat{r}), \hat{\varepsilon}_{t-1} 1(|\hat{\varepsilon}_{t-1}| < \hat{r})). \end{aligned}$$

La consistencia de $\hat{V}(\hat{\beta})$ se encuentra en la prueba del Teorema 3.5. Por lo tanto, bajo la hipótesis nula y usando el resultado de este Teorema, el estadístico de contraste convergerá asintóticamente a una χ^2_1 . Como en el caso Observable, un supuesto implícito para que las perturbaciones pequeñas sean transitorias es I.1.

Contraste para la existencia de Threshold cuando r^0 es desconocido

En esta sección analizaremos el contraste de existencia de efecto threshold o contraste de simetría cuando el valor de r^0 es desconocido y distinto de 0. Como ya hemos comentado, en este contexto y bajo la hipótesis nula, el parámetro r no estará identificado y no habrá superconsistencia de los estimadores, lo que invalida la teoría anterior. Como en el caso Observable, planteamos la siguiente regresión:

$$x_t = \mu + \rho x_{t-1} - \theta \varepsilon_{t-1} - \alpha(r) \varepsilon_{t-1} 1(g(\varepsilon_{t-1}) < r) + \varepsilon_t$$

Basada en esta regresión, la hipótesis nula de no existencia de efecto threshold será:

$$H_0 : \alpha(r) = 0$$

y la alternativa,

$$H_a : \alpha(r) \neq 0,$$

para $r \in [\underline{r}, \bar{r}]$. Al igual que en el caso Observable, proponemos el $\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r)$ como estadístico de contraste, con

$$\Gamma_T(r) = \frac{T \hat{\alpha}^2(r)}{\hat{V}(\hat{\alpha}(r))},$$

donde

$$\begin{aligned} \hat{V}(\hat{\alpha}(r)) &= \hat{\sigma}_\varepsilon^2 R' H_T^{-1}(\hat{\beta}, r) R \\ R' &= (0, 1, 0, 0) \\ H_T^{-1}(\hat{\beta}, r) &= T^{-1} \sum_{t=1}^T h_t(\hat{\beta}, r) h_t'(\hat{\beta}, r) \\ h_t'(\hat{\beta}, r) &= \left(\sum_{j=1}^t \theta^{0j-1} e_{t-j}(\hat{\beta}, r), \sum_{j=1}^t \theta^{0j-1} e_{t-j}^-(\hat{\beta}, r), -\sum_{j=1}^t \theta^{0j-1} x_{t-1}, -\sum_{j=1}^t \theta^{0j-1} \right) \\ e_{t-j}^-(\hat{\beta}, r) &= 1 \left(g(e_{t-j}(\hat{\beta}, r)) < r \right) e_{t-j}(\hat{\beta}, r). \end{aligned}$$

La distribución asintótica del estadístico vendrá determinada por los siguientes procesos, $s(r)$, que será un proceso gaussiano de media cero y función de covarianzas igual a $K(r_1, r_2) = \sigma_\varepsilon^2 E(h_t(\beta^0, r_1) h_t'(\beta^0, r_2))$. También por $\bar{S}(r) = R' H^{-1}(r) s(r)$, que de nuevo será un proceso gaussiano de media cero y función de covarianzas $\bar{K}(r_1, r_2) = R' H^{-1}(r_1) K(r_1, r_2) H^{-1}(r_2) R$, con

$$H(r) = E(h_t(\beta^0, r) h_t'(\beta^0, r)).$$

y R descrita anteriormente.

El siguiente Teorema muestra la distribución asintótica del estadístico propuesto en esta sección, donde \Rightarrow implica convergencia uniforme en distribución.

Teorema 3.6 *Bajo H_0 y los supuestos G.0, G.2, G.3, y ASZ.1-ASZ.2, en el caso del tamaño, o ASG.1-ASG.2 en el caso del signo, para cualquier funcional continuo $G(\cdot) : [\underline{r}, \bar{r}] \rightarrow R$, se tiene que*

$$G(\Gamma_T(r)) \Rightarrow G\left(\overline{S}'(r) \overline{K}^{-1}(r, r) \overline{S}(r)\right),$$

Prueba, ver Anexo 2.

Dado que $\sup_{r \in [\underline{r}, \bar{r}]}$ es un funcional continuo de $[\underline{r}, \bar{r}] \rightarrow R$, tenemos que bajo H_0 :

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{T\hat{\alpha}^2(r)}{\widehat{V}(\hat{\alpha}(r))} \Rightarrow \sup_{r \in [\underline{r}, \bar{r}]} \overline{S}'(r) \overline{K}^{-1}(r, r) \overline{S}(r).$$

Exceptuando la matriz de varianzas y covarianzas del proceso, en su definición de z_t , la distribución asintótica del estadístico es la misma que en el caso Observable, y por tanto, para un r fijo esta coincidirá con una χ_1^2 , ver el caso de $r^0 = 0$ conocido. Igualmente, la distribución bajo H_0 de $G(\Gamma_T(r))$ dependerá, en general, de la función de covarianzas $\overline{K}(r_1, r_2)$, y los valores críticos no podrán ser tabulados excepto en casos especiales. Por lo tanto, para poder realizar el contraste proponemos la misma aproximación bootstrap que describimos en el capítulo anterior para el caso Observable.

Cuando la hipótesis nula de no efecto threshold sea rechazada, podemos pasar a contrastar la hipótesis de existencia de perturbaciones transitorias en el caso del tamaño y para las perturbaciones pequeñas, utilizando el estadístico descrito anteriormente a través de la estimación en dos etapas.

Finalmente, el resultado del Teorema permite extender el análisis a cualquier otro estadístico basado en un funcional continuo de $\Gamma_T(r)$.

3.4.3. Comportamiento de los contrastes en muestras finitas

Como hicimos en el capítulo anterior, en esta sección se lleva a cabo un breve estudio sobre el comportamiento de los contrastes para la existencia de threshold, propuestos y

desarrollados en las secciones anteriores. Con la intención de realizar un análisis comparativo, los parámetros del estudio de Monte Carlo serán los mismos que los seleccionados en el Capítulo anterior. Por tanto, el tamaño del Monte Carlo será $M=400$ y el del Wild Bootstrap $B=400$. Igualmente, el estudio se realizará para los niveles de significación $\alpha = 0.05, 0.10$ y los tamaños muestrales $T = 200, 600$.

En primer lugar, analizaremos el comportamiento del contraste de existencia de threshold para el caso en que la variable threshold es el signo de la perturbación, descrito en la Sección 3.4.1. Para el estudio del tamaño generaremos las mismas series ARIMA(0,1,1) del capítulo anterior, esto es,

$$(1 - L)y_t = x_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

con $\varepsilon_t \text{ iid } N(0,1)$ y $\theta = (-0.5, 0, 0.5)$. La tabla 3.1 muestra los resultados, que son muy similares al caso Observable tratado en el Capítulo anterior, aunque conviene destacar un mejor comportamiento de la aproximación asintótica en este caso.

Para el estudio de la potencia generaremos el siguiente proceso ARTIMA-Shock,

$$(1 - L)y_t = x_t = \varepsilon_t - \theta_1\varepsilon_{t-1}1(\varepsilon_{t-1} > 0) - \theta_2\varepsilon_{t-1}1(\varepsilon_{t-1} < 0)$$

para los valores de los parámetros $\theta_1 = 0$ y $\theta_2 = (-0.8, -0.6, -0.4, -0.2, 0.2, 0.4, 0.6, 0.8)$ donde ε_t es $\text{iid } N(0,1)$. Las tablas 3.3 y 3.4 muestran los resultados del estudio de potencia para $T = 200$ y $T = 600$ respectivamente.

Los resultados son similares a los obtenidos en el caso Observable. De nuevo, a medida que aumenta $\partial = |\theta_1 - \theta_2|$, aumenta la potencia. En este caso, al ser el porcentaje y varianza de las perturbaciones negativas constante para los distintos valores de θ_2 , la potencia es ligeramente mayor para los valores de $\theta_2 > 0$.

Para el caso en que r^0 es desconocido, y por lo tanto, en el análisis del comportamiento del estadístico $\sup_{r \in [\underline{r}, \bar{r}]} \Gamma_T(r)$, nos centraremos en el caso del ARTIMA-Shock-Size, donde es el tamaño de la perturbación, $|\varepsilon_t|$, la variable threshold. Como en los casos anteriores,

para el estudio del tamaño generamos las series ARIMA(0,1,1). Los resultados se muestran en la Tabla 3.2, y como sucedía en el caso observable su comportamiento es superior al caso del signo.

Finalmente, para la potencia generamos el proceso ARTIMA-Shock-Size,

$$(1 - L)y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| > r) - \theta_2 \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r)$$

con $(\theta_1, \theta_2) = (0, 1)$, $r = (0.35, 0.5, 0.7)$ y $\varepsilon_t \text{ iid } N(0, 1)$, esto es, el caso de perturbaciones pequeñas transitorias. Los resultados se encuentran en la Tabla 3.5. Como vimos en la Sección 3.2.2 a través de las Figuras 3.7 y 3.8, un porcentaje de perturbaciones transitorias inferior al 50 % garantiza la invertibilidad del proceso. Para la muestra usada en este estudio, ese valor se obtiene para el caso $r = 0.7$. Igualmente, la consistencia se garantiza para un porcentaje de transitorios entorno al 30 %, que coincidiría en nuestra muestra con el caso $r = 0.5$.

Como en el caso Observable tratado en el Capítulo 2, de nuevo, cuanto mayor es el valor de r , mayor es la potencia del contraste, debido a que fijado (θ_1, θ_2) , con r también aumenta el porcentaje y varianza de las perturbaciones transitorias. Sin embargo, el contraste de existencia de threshold en el ARTIMA-Shock presenta una potencia sensiblemente inferior. La razón se encuentra, fundamentalmente, en que la varianza de las perturbaciones en el régimen “transitorio” es inferior en el caso del ARTIMA-Shock que en el ARTIMA-Observable.

Finalmente, aunque la potencia para los casos en los que el modelo se puede estimar de forma consistente, $r < 0.5$, pueda parecer baja, en el Apéndice 1 de este Capítulo se muestra que el comportamiento del contraste propuesto es netamente superior al contraste de incorrelación en diversos contextos lineales.

Cuadro 3.1: Tamaño del contraste de existencia de Threshold. $r^0 = 0$

θ	Valores Críticos Bootstrap				Valores Críticos Asintóticos			
	T=200		T=600		T=200		T=600	
	$\alpha - level$				$\alpha - level$			
	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
-0.5	0.052	0.090	0.037	0.082	0.055	0.100	0.042	0.077
0.0	0.047	0.090	0.042	0.077	0.055	0.105	0.042	0.075
0.5	0.070	0.125	0.080	0.135	0.092	0.150	0.090	0.137

Cuadro 3.2: Tamaño del contraste de existencia de Threshold. r^0 desconocido.

	T=200		T=600	
	$\alpha - level$			
θ	0.05	0.10	0.05	0.10
-0.5	0.052	0.097	0.0325	0.080
0.0	0.050	0.112	0.0425	0.105
0.5	0.055	0.107	0.0475	0.087

3.5. Aplicación al Índice de Precios Industriales

En esta sección aplicaremos los resultados obtenidos para el modelo ARTIMA-Shock-Sign al índice de precios industriales (PPI), con la intención de comparar los resultados obtenidos por Wecker (1981) con los obtenidos a través de la metodología propuesta en este capítulo. Además se extenderá hasta la actualidad el análisis de una de las series, en la que parece haber un cambio estructural que modifica el comportamiento de los precios.

Wecker (1981) ajusta los modelos de medias móviles asimétricos en el signo de la per-

Cuadro 3.3: Potencia del contraste de existencia de Threshold, $r^0 = 0$.

$\theta_1 = 0$	θ_2							
	-0.8	-0.6	-0.4	-0.2	0.2	0.4	0.6	0.8
$\alpha - level$	Valor Crítico Bootstrap							
0.05	0.970	0.790	0.480	0.145	0.130	0.490	0.825	0.980
0.10	0.995	0.885	0.605	0.235	0.215	0.615	0.900	1.000
$\alpha - level$	Valor Crítico Asintótico							
0.05	0.955	0.755	0.445	0.170	0.140	0.430	0.770	0.955
0.10	0.985	0.835	0.570	0.245	0.205	0.530	0.845	0.990

Cuadro 3.4: Potencia del contraste de existencia de Threshold, $r^0 = 0$.

$\theta_1 = 0$	θ_2							
	-0.8	-0.6	-0.4	-0.2	0.2	0.4	0.6	0.8
$\alpha - level$	Valor Crítico Bootstrap							
0.05	1.000	1.000	0.875	0.350	0.265	0.875	1.000	1.000
0.10	1.000	1.000	0.915	0.465	0.370	0.915	1.000	1.000
$\alpha - level$	Valor Crítico Asintótico							
0.05	1.000	0.995	0.860	0.370	0.275	0.860	1.000	1.000
0.10	1.000	1.000	0.905	0.480	0.370	0.905	1.000	1.000

En este caso, para todos los valores de θ_2 el porcentaje y varianza media de las perturbaciones negativas es la misma, para $T = 200$, $P(\varepsilon_t < 0) = 0.501$ y $Var(\varepsilon_t 1(\varepsilon_t < 0)) = 0.378$, y para $T = 600$, $P(\varepsilon_t < 0) = 0.500$ y $Var(\varepsilon_t 1(\varepsilon_t < 0)) = 0.352$.

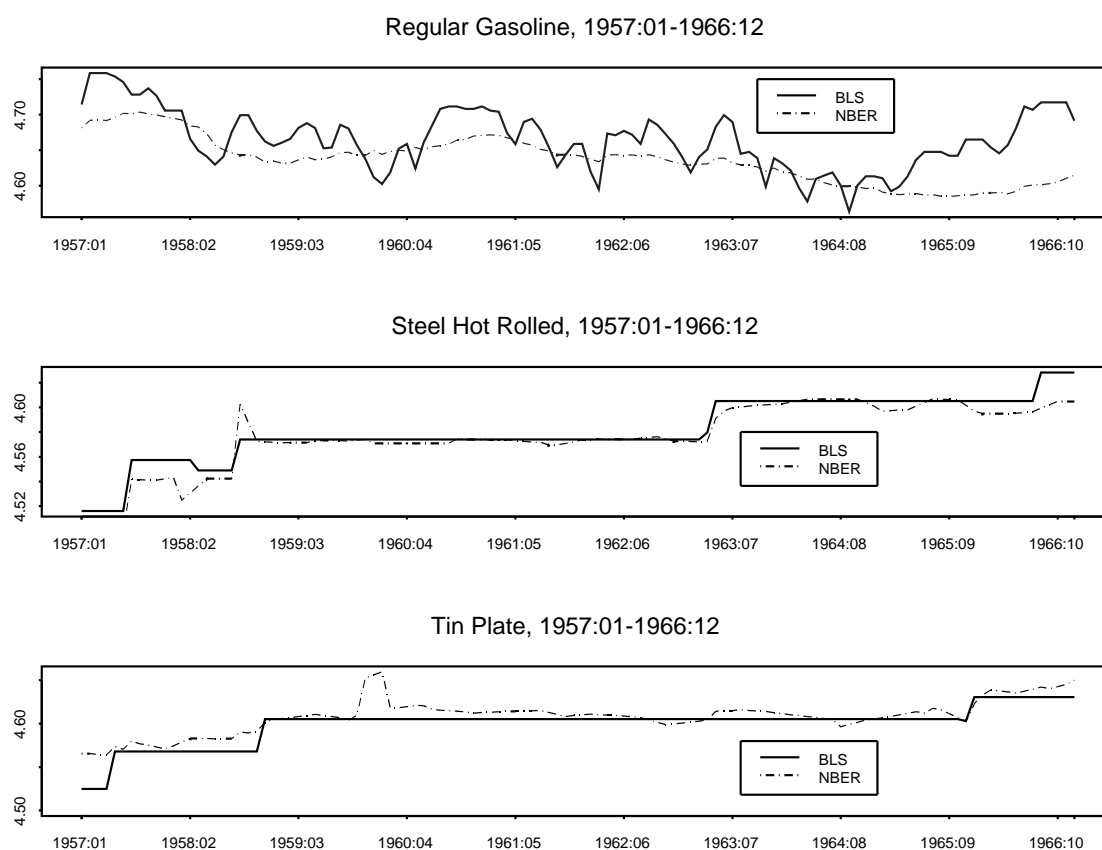
Cuadro 3.5: Potencia del contraste de existencia de Threshold. r^0 desconocido.

$\theta_1 = 0, \theta_2 = 1$	T=200			T=600		
	r					
	$\alpha - level$	0.35	0.5	0.7	0.35	0.5
0.05	0.190	0.335	0.860	0.390	0.910	1.000
0.10	0.325	0.565	0.935	0.555	0.980	1.000
$P(\varepsilon_t < r)$	0.269	0.377	0.511	0.273	0.382	0.516
$\text{Var}(\varepsilon_t 1(\varepsilon_t < r))$	0.010	0.030	0.078	0.011	0.030	0.079

turbación a seis series de precios industriales seleccionadas entre las diferentes series recopiladas por Stigler y Kindahl (1970). Las series seleccionadas fueron las de los precios del estaño, gasolina y acero, que denominaremos por Tin Plate, Regular Gasoline y Steel Hot Rolled. Cada precio fue medido basandose en dos métodos diferentes, uno propuesto por el Bureau of Labor Statistic (BLS) y el otro, por los autores y denotado por NBER. El primero se centra en los precios 'spot' y 'quoted', el último, en los precios de contrato. El gráfico del logaritmo de cada serie puede encontrarse en la figura 3.9. Stigler y Kindahl encuentran que el índice BLS crece respecto al índice NBER entorno a 0.7, y establece que la única explicación plausible es una inercia asimétrica en el movimiento de los precios industriales. Siguiendo su argumentación, esta asimetría vendría explicada por la prudencia de los vendedores a la hora de reducir los precios, y el comportamiento opuesto para incrementar el precio. Además, en el contexto de inflación positiva, la reducción de precios revierte más a menudo que su incremento. Todo esto implica que las perturbaciones negativas serán menos persistentes que las positivas. Claramente, este comportamiento es consistente con el precio del vendedor, el cual está mejor recogido en el precio 'spot' o 'quoted', y por lo tanto, en el índice BLS, pero no en el precio del comprador, medido mejor por el precio de

contrato, y por lo tanto, en el índice NBER. Según este razonamiento, es más probable que el índice BLS presente asimetrías en la persistencia basada en el signo de la perturbación que el índice NBER.

Figura 3.9: Índice de Precios Industriales, 1957:01-1966:12



En la Tabla 3.6 se encuentran los resultados de estimación por CLS para los modelos ARIMA y ARTIMA-Shock-Sign. La especificación de cada modelo es la escogida por Wecker. Para realizar el contraste de asimetría en persistencia que verifique la validez de la argumentación expuesta por Stigler y Kindahl, usaremos el estadístico descrito en la

Sección 3.4.1:

$$\frac{T \left(R' \hat{\beta} \right)^2}{\left[\hat{\sigma}_\varepsilon^2 \left(R' \hat{H}^{-1} R \right) \right]}$$

con $R' = (1, -1, 0, 0)$, el cual bajo la hipótesis nula de no existencia de efecto threshold convergerá a una χ_1^2 , por el resultado del Teorema 3.3. En su artículo, Wecker propone el estimador Máximo Verosímil y un estadístico basado en la razón de verosimilitudes. Como puede verse en la Tabla 3.6, los resultados de estimación y contraste obtenidos con los métodos propuestos en este Capítulo son muy similares a los obtenidos por Wecker, esto es, las tres series BLS presentan asimetrías en la persistencia basadas en el signo de la perturbación, mientras que esta hipótesis es rechazada para las tres series NBER.

Ahora, dado que el índice BLS para los precios de la gasolina extendido hasta el 2004 esta disponible, completamos el análisis anterior para este precio con este periodo. En la Figura 3.10 se representa el logaritmo de la serie desde 1957:01 a 2004:02. La serie muestra un cambio estructural durante los años 70 debido a la crisis del petróleo. A continuación analizaremos si esta crisis afecta al comportamiento de los precios, concretamente, al comportamiento asimétrico de las perturbaciones. Para ello, ajustaremos el modelo ARTIMA-Shock-Sign al índice de precios en dos submuestras, la primera, anterior a la crisis del petróleo, periodo 1957:01-1970:01, y la segunda, posterior a la crisis, periodo 1983:01-2004:02. También presentamos los resultados para el ajuste de toda la muestra. Los resultados se muestran en la Tabla 3.7.

Para el periodo anterior a la crisis, ajustamos un ARIMA(2,1,1) para el caso lineal, dado que los residuos del ARIMA(0,1,1) propuesto por Wecker presentan cierta estructura de correlación. Por la misma razón, para el caso no lineal, ajustamos un ARTIMA(1,1,1,1). El estadístico para el contraste de asimetría es igual a 5.043, por lo tanto rechazamos la hipótesis nula de persistencia simétrica de las perturbaciones en función de su signo al 5 %. Sin embargo, esta conclusión cambia para el periodo posterior a la crisis. En este periodo, no

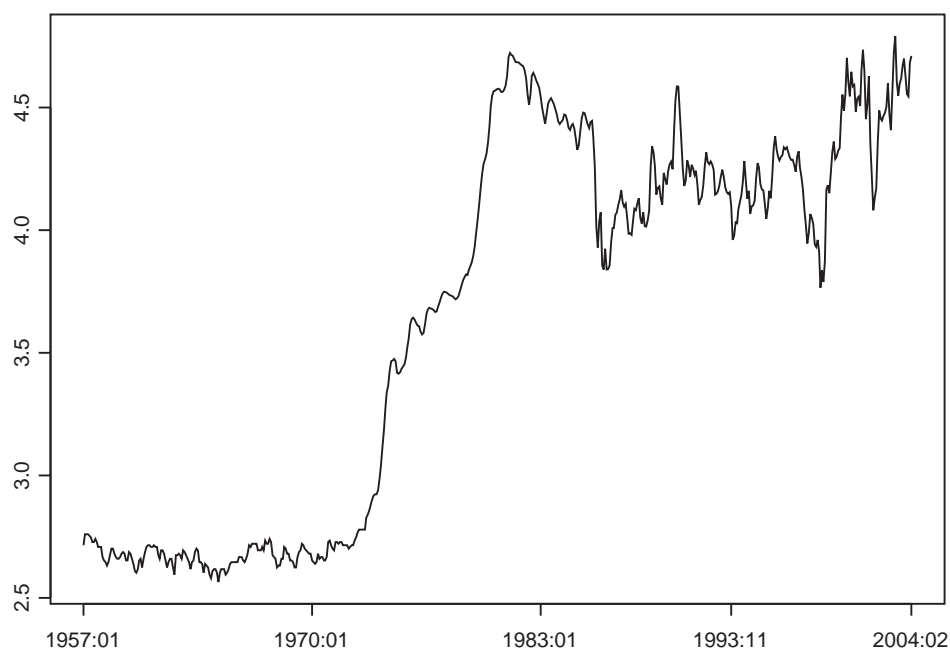
Cuadro 3.6: Resultados de Estimación y Contraste para los modelos ARIMA y ARTIMA-Shock-Sign para los índices PPI

	ARIMA			ARTIMA					$\chi^2_{1[0,05]}$
	θ	$\sigma_\varepsilon \times 10^2$	AIK	μ	θ_1	θ_2	$\sigma_\varepsilon \times 10^2$	AIK	3,84
S. H. Rol.									
BLS	-0.040 (0.091)	0.552	-7.54	0.000 (0.000)	0.040 (0.096)	0.920 (0.103)	0.532	-7.58	6.08
NBER	0.045 (0.091)	0.764	-6.90	0.001 (0.001)	0.094 (0.099)	-0.393 (0.298)	0.744	-6.91	1.45
T. Plate									
BLS	0.015 (0.092)	0.582	-7.44	0.000 (0.000)	0.012 (0.092)	1.046 (0.088)	0.555	-7.50	8.07
NBER	-0.097 (0.091)	0.653	-7.21	0.000 (0.000)	-0.169 (0.132)	0.016 (0.146)	0.643	-7.20	0.89
R. Gaso.									
BLS	-0.205 (0.090)	2.016	-4.96	-0.008 (0.002)	-0.459 (0.149)	0.491 (0.174)	1.928	-5.00	11.4
NBER	0.633 (0.134)	0.361	-8.39	-0.000 (0.000)	-0.251 (0.156)	-0.311 (0.160)	0.376	-8.269	0.05

En paréntesis las desviaciones estándar correspondientes. En los modelos lineales, la constante esta restringida a ser 0. Ecuación del modelo ARTIMA: $x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1}^+ - \theta_2 \varepsilon_{t-1}^-$.

se puede rechazar la hipótesis de simetría en la persistencia. Además, bajo la modelización simétrica o lineal, las perturbaciones son claramente más persistentes en este periodo más reciente. Cabe destacar también la mayor varianza de las perturbaciones en ambos modelos, lineal y no lineal, en el periodo posterior a la crisis. Esta mayor varianza esta relacionada con el rechazo de la asimetría en este periodo. Como se comento anteriormente, la hipótesis de asimetría en la persistencia de las perturbaciones se basa en la oposición del vendedor a la variación en los precios, sobre todo cuando esta variación implica una reducción del

Figura 3.10: Precio Regular Gasoline, Índice BLS, 1957:01-2004:02



mismo. Una mayor varianza en las perturbaciones implica una menor reticencia al cambio en los precios, y por tanto, probablemente a su disminución. A la vista de estos resultados, podemos concluir que la crisis del petróleo provocó un cambio en el comportamiento de los precios de la gasolina, y no sólo un mero incremento de precios a través de un cambio en la constante.

3.6. Conclusión

En este Capítulo estudiamos en detalle los modelos ARTIMA-Shock, que se caracterizan porque la asimetría en la persistencia es provocada por alguna propiedad de la propia perturbación. En concreto, se analizan las propiedades del signo y tamaño. Como en el caso

Cuadro 3.7: Resultados de estimación de los modelos ARIMA y ARTIMA-Shock-Sign para el índice BLS de precios de la Regular Gasoline

	ARIMA(2,1,1)	ARTIMA(1,1,1,1)	ARIMA(1,1,1)	ARTIMA(1,1,1,1)	ARIMA(1,1,1)	ARTIMA(1,1,1,1)
	1957:01-1970:01		1983:01-2004:02		1957:01-2004:02	
μ	-0.000 (0.000)	-0.003 (0.001)	0.000 (0.005)	-0.003 (0.007)	0.003 (0.002)	0.002 (0.003)
ρ_1	0.985 (0.080)	0.577 (0.150)	-0.392 (0.112)	-0.391 (0.101)	-0.364 (0.076)	-0.371 (0.070)
ρ_2	-0.163 (0.079)					
θ	0.983 (0.008)		-0.757 (0.080)		-0.743 (0.054)	
θ_1		0.508 (0.194)		-0.878 (0.090)		-0.850 (0.065)
θ_2		0.929 (0.131)		-0.685 (0.094)		-0.686 (0.064)
$\sigma_\varepsilon \times 10^2$	1.974	2.005	6.673	6.602	4.759	4.722
AIK	-4.961	-4.929	-2.564	-2.566	-3.246	-3.253
$\chi^2_{1[0,05]} = 3,84$	5.043		2.272		3.474	

Entre paréntesis la desviación estándar correspondiente.

Observable, se obtienen las condiciones de invertibilidad del modelo, en las que interviene el grado de discontinuidad del mismo. Estas condiciones permiten que las perturbaciones pequeñas puedan ser transitorias. Este resultado no es extensible al caso del signo, donde la condición I.1 impide que pueda haber perturbaciones transitorias aunque exista un régimen no invertible en la media móvil.

El método de estimación propuesto será el de mínimos cuadrados condicionales. Para

el caso del signo con $r^0 = 0$ conocido, los resultados son similares a los obtenidos en el caso Observable, consistencia y normalidad asintótica de los estimadores. Este resultado supone una aportación fundamental del capítulo, ya que a pesar de la amplia utilización de este modelo, no se habían obtenido, hasta ahora, resultados de validez asintótica de los métodos propuestos. Para el resto de casos, $|r^0| > 0$, los resultados son diferentes a los obtenidos hasta ahora, probándose la superconsistencia de todos los estimadores. Esto hace, que para obtener la normalidad asintótica de los mismos sea necesario utilizar estimaciones en dos etapas.

Respecto al contraste de existencia de threshold, para el caso del signo con $r^0 = 0$ se propone un contraste de Wald (ya propuesto por Brännäs y De Gooijer, 1994), y cuya distribución bajo H_0 será una χ_1^2 . Para el otro caso, $|r^0| > 0$, el contraste usa un estadístico basado en el supremo similar al del caso Observable, del que conocemos su distribución asintótica, aunque no es tabulable. Por lo tanto, y como en el capítulo anterior, se propone una aproximación Wild Bootstrap para obtener el p – *valor* del contraste. Un pequeño estudio de Monte Carlo corrobora el buen comportamiento de los contrastes propuestos. Además, cuando se rechaza la hipótesis nula de simetría, para el caso particular de los Modelos ARTIMA-Shock-Size, se propone el contraste de existencia de perturbaciones pequeñas transitorias. Este se traduce en el contraste de existencia de raíz unitaria en el régimen MA correspondiente a las perturbaciones pequeñas, que de nuevo se basa en un estadístico de Wald cuya distribución asintótica es una χ_1^2 .

Finalmente, el procedimiento propuesto se aplica a los diferentes índices de precios usados por Wecker, obteniéndose similares resultados, esto es, la existencia de asimetrías en el signo en los índices BLS. Además, el análisis se extiende, en el caso del precio de la gasolina, al periodo posterior a la crisis del petróleo, probándose que esta no solo provoca un cambio en la deriva, sino también en su dinámica, de la que desaparecen las asimetrías provocadas por el signo.

3.7. Bibliografía

Brännäs, K. y De Gooijer, J.G. (1994): “Autoregressive – asymmetric Moving Average models for Business Cycle data”. *Journal of Forecasting*, Vol. **13**, 529-544.

Brännäs, K., De Gooijer, J.G. y Teräsvirta, T. (1998): “Testing linearity against nonlinear moving average models”. *Commun. Statist. - Theory Meth.*, Vol. **27**, 2025-2035.

Brännäs, K y de Luna, X. (1998): “Generalized Method of Moment and Indirect Estimation of Autoregressive – Asymmetric Moving Average Models”. *Computational Statistics*, Vol. **13**, 485-494.

Elwood, S.K. (1998): “Is the persistence of shocks to output asymmetric?”. *Journal of Monetary Economics*, Vol. **41**, 411-426.

Engle, R.F. y Smith, A.D. (1999): “Stochastic Permanent Breaks”. *The Review of Economics and Statistics*, Vol. **81**, 553-574.

Guay, A. y Scaillet, O. (2003): “Indirect Inference, Nuisance Parameter, and Threshold Moving Average Models”. *Journal of Business and Economic Statistics*, Vol. **21**, 122-32.

Hamilton, J. (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle”. *Econometrica* Vol. **75**, N2, 357-384.

Stigler, G.J. y Kindahl, J.K. (1970): *The Behavior of Industrial Prices*. National Bureau of Economic Research.

Wecker, W.E. (1981): “Asymmetric time series”. *Journal of the American Statistical Association*, Vol. **76**, N 373, 16-21.

3.8. Anexo 1.

En este Anexo se lleva a cabo un estudio comparativo entre el contraste de existencia de efecto threshold provocado por el tamaño de la perturbación y el contraste de incorrelación en el contexto lineal. Para ello consideraremos dos aproximaciones lineales al modelo ARTIMA-Shock-Size:

$$(1 - L) y_t = \varepsilon_t - \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r) = \varepsilon_t^L + \varepsilon_t^s - \varepsilon_{t-1}^s$$

utilizado en la Sección 3.4.3. La primera corresponde a su representación de Wold:

$$(1 - L)y_t = u_t - \theta u_{t-1}$$

tal que

$$\begin{aligned} (1 + \theta^2) \text{Var}(u_t) &= \text{Var}(\varepsilon_t) + \text{Var}(\varepsilon_t^s) \\ \theta \text{Var}(u_t) &= \text{Var}(\varepsilon_t^s) \end{aligned}$$

donde u_t son *iid* $N(0, \sigma_u)$ con $\sigma_u > 0$. La segunda, será la representación lineal de componentes inobservados incorrelados (UC-0) (ver Harvey, 1985 y Clark, 1987):

$$(1 - L) y_t = z_t + u_t - u_{t-1}$$

con,

$$\begin{aligned} \text{Var}(z_t) &= \text{Var}(\varepsilon_t^L) \\ \text{Var}(u_t) &= \text{Var}(\varepsilon_t^s) \end{aligned}$$

donde u_t es *iid* $N(0, \sigma_u)$ y z_t es *iid* $N(0, \sigma_z)$ e independientes entre si, con $\sigma_u > 0$ y $\sigma_z > 0$.

Para ambos modelos lineales compararemos el contraste de incorrelación de $(1 - L)y_t$, esto es, el contraste de la hipótesis nula $H_0 : \theta = 0$ en la representación de Wold y $H_0 : \text{Var}(u_t) = 0$ en el modelo UC-0. Ambas hipótesis coinciden con la hipótesis de

ausencia de efecto threshold en el modelo ARTIMA. Para el contraste de incorrelación en ambos modelos lineales, ajustaremos a las series el siguiente modelo ARIMA(0,1,1):

$$(1 - L)y_t = \eta_t + \theta\eta_{t-1}$$

contrastando la hipótesis nula $H_0 : \theta = 0$ a través del estadístico:

$$WL = \frac{\hat{\theta}^2}{\hat{\sigma}_{\hat{\theta}}^2}.$$

Bajo las propiedades de u_t y z_t , y dado que $\theta \in (-1, 1)$, la distribución asintótica de WL bajo H_0 será una χ^2 de un grado de libertad. Los resultados se muestran en la Tabla 3.8. Como en el caso del estudio realizado en la Sección 3.4.3, este se realiza para los tamaños muestrales $T = 200, 600$, y los niveles de significación $\alpha = 0.05, 0.10$.

En el primer bloque de la tabla se reproducen los resultados para el contraste threshold analizado en la Sección 3.4.3, para el caso del tamaño de las perturbaciones y los distintos valores de r considerados. En el segundo bloque, se muestran los resultados del estudio para el tamaño y potencia del contraste de incorrelación cuando la aproximación al modelo ARTIMA-Shock-Size se realiza a través de su representación de Wold. También se reproducen los valores de θ y la $Var(u_t)$ correspondientes a cada r . Por último, en el tercer bloque, se presentan los resultados de este contraste cuando la aproximación se realiza a través del modelo UC-0. Igualmente se muestran los valores de la $Var(u_t)$ y $Var(z_t)$ correspondientes a cada r . Como ya se comentó en dicha sección y prueban los resultados de la Tabla 3.8, el comportamiento del contraste de existencia de threshold es netamente superior al de incorrelación en ambos contextos lineales cuando la desviación de la nula es pequeña. Esta diferencia es superior para el caso de la aproximación de Wold, caso que prácticamente no presenta potencia para los distintos valores de r considerados, exceptuando $r = 0.7$.

Cuadro 3.8: Comparación del contraste threshold para el modelo ARTIMA y de incorrelación en el modelo lineal

		T=200				T=600			
		0	0.35	0.50	0.70	0	0.35	0.50	0.70
ARTIMA	r								
	$\text{Var}(\varepsilon_t^s)$	0.000	0.010	0.030	0.078	0.000	0.011	0.030	0.079
	α								
	0.05	0.050	0.190	0.335	0.860	0.0425	0.390	0.910	1.000
	0.10	0.112	0.325	0.565	0.935	0.105	0.555	0.980	1.000
ARIMA (R. Wold)	θ	0.000	0.006	0.019	0.050	0.000	0.006	0.019	0.050
	$\text{Var}(u_t)$	1.000	1.565	1.565	1.562	1.000	1.565	1.565	1.562
	α								
	0.05	0.070	0.070	0.085	0.145	0.035	0.037	0.057	0.262
	0.10	0.110	0.117	0.135	0.217	0.087	0.090	0.130	0.390
UC-0	$\text{Var}(z_t)$	1.000	1.032	1.013	0.965	1.000	1.032	1.013	0.965
	$\text{Var}(u_t)$	0.000	0.010	0.030	0.078	0.000	0.010	0.030	0.078
	α								
	0.05	0.065	0.062	0.087	0.210	0.055	0.057	0.100	0.425
	0.10	0.105	0.110	0.145	0.342	0.097	0.135	0.162	0.525

3.9. Anexo 2.

Prueba del Lema 3.1.

$$\Phi_p(L) x_t = \mu + \varepsilon_t - \theta_1 1(|\varepsilon_{t-1}| > r) \varepsilon_{t-1} - \theta_2 1(|\varepsilon_{t-1}| < r) \varepsilon_{t-1},$$

por lo tanto,

$$\widehat{\varepsilon}_t = \Phi_p(L) x_t - \mu + \theta_1 1(|\widehat{\varepsilon}_{t-1}| > r) \widehat{\varepsilon}_{t-1} + \theta_2 1(|\widehat{\varepsilon}_{t-1}| < r) \widehat{\varepsilon}_{t-1}.$$

Definimos $\partial = |\theta_1 - \theta_2|$ y los sucesos

$$A_{1,t-1} = \{|\varepsilon_{t-1}| > r, |\widehat{\varepsilon}_{t-1}| > r\} \quad A_{2,t-1} = \{|\varepsilon_{t-1}| < r, |\widehat{\varepsilon}_{t-1}| < r\} \quad (3.3)$$

$$A_{3,t-1} = \{|\varepsilon_{t-1}| > r, |\widehat{\varepsilon}_{t-1}| < r\} \quad A_{4,t-1} = \{|\varepsilon_{t-1}| < r, |\widehat{\varepsilon}_{t-1}| > r\}.$$

Entonces, $v_t = \widehat{\varepsilon}_t - \varepsilon_t$, se puede escribir como,

$$\begin{aligned} v_t &= \theta_1 1(A_{1,t-1}) v_{t-1} + \theta_2 1(A_{2,t-1}) v_{t-1} + \theta_1 1(A_{3,t-1}) v_{t-1} + (\theta_2 - \theta_1) 1(A_{3,t-1}) \widehat{\varepsilon}_{t-1} + \\ &\quad + \theta_1 1(A_{4,t-1}) v_{t-1} + (\theta_1 - \theta_2) 1(A_{4,t-1}) \varepsilon_{t-1} \\ &= (\theta_1 - \theta_2) [1(A_{4,t-1}) \varepsilon_{t-1} - 1(A_{3,t-1}) \widehat{\varepsilon}_{t-1}] + [\theta_1 1(A_{2,t-1}^c) + \theta_2 1(A_{2,t-1})] v_{t-1}, \end{aligned} \quad (3.4)$$

con $A_{2,t-1}^c$ el suceso complementario de $A_{2,t-1}$. Si $|\theta_1| < |\theta_2|$, tenemos que

$$|v_t| \leq \partial [1(A_{4,t-1}) |\varepsilon_{t-1}| + 1(A_{3,t-1}) |\widehat{\varepsilon}_{t-1}|] + [|\theta_1| 1(|\varepsilon_{t-1}| > r) + |\theta_2| 1(|\varepsilon_{t-1}| < r)] |v_{t-1}|,$$

con $\partial = |\theta_1 - \theta_2|$, y por tanto,

$$E|v_t| \leq \partial r E[1(A_{4,t-1}) + 1(A_{3,t-1})] + E\{[|\theta_1| 1(|\varepsilon_{t-1}| > r) + |\theta_2| 1(|\varepsilon_{t-1}| < r)] |v_{t-1}|\}.$$

Usando que v_{t-1} es \mathfrak{F}_{t-2} medible y $f^* = \max_e 2f(e)$,

$$E[1(A_{4,t-1}) + 1(A_{3,t-1})] \leq f^* E(|v_{t-1}|). \quad (3.5)$$

Por lo tanto, tomando $\lambda_1^* = [\partial r f^* + E(|\theta(\varepsilon_{t-1})|)]$,

$$E|v_t| \leq [\partial r f^* + E(|\theta(\varepsilon_{t-1})|)] E|v_{t-1}| = \lambda_1^{*,t} E|v_0|.$$

Sin embargo, para la invertibilidad será necesario la convergencia de $E(v_{t-1}^2)$, entonces

$$\begin{aligned} v_t^2 &\leq \partial^2 r^2 [1(A_{4,t-1}) + 1(A_{3,t-1})] + [\theta_1^2 1(|\varepsilon_{t-1}| > r) + \theta_2^2 1(|\varepsilon_{t-1}| < r)] v_{t-1}^2 + \\ &\quad + 2\partial r [1(A_{4,t-1}) + 1(A_{3,t-1})] [|\theta_1| 1(|\varepsilon_{t-1}| > r) + |\theta_2| 1(|\varepsilon_{t-1}| < r)] |v_{t-1}|. \end{aligned}$$

Definiendo $K = \partial^2 r^2 f^* + 4\partial r$,

$$E(v_t^2) \leq K E|v_{t-1}| + E(|\theta(\varepsilon_{t-1})|) E(v_{t-1}^2) \leq K \lambda_1^{*,t-1} E|v_0| + \lambda_1^* E(v_{t-1}^2) \leq t K \lambda_1^{*,t-1} E|v_0|.$$

Dado que por el supuesto ASZ.0 a), $\lambda_1^* < 1$, se cumple que $\lim_{t \rightarrow \infty} E(v_t^2) = 0$, y por lo tanto, el proceso será invertible.

Por otro lado, si $|\theta_2| < |\theta_1|$, de la ecuación (3.4) y usando ASZ.0 b), tenemos que

$$\begin{aligned} |v_t| &\leq \partial [1(A_{4,t-1}) |\varepsilon_{t-1}| + 1(A_{3,t-1}) |\widehat{\varepsilon}_{t-1}|] + |\theta_1| \\ E|v_t| &\leq \lambda_1^{*,t} E|v_0|, \end{aligned}$$

con $\lambda_1^* < 1$. El resto de la prueba sigue como en el caso anterior, $|\theta_1| < |\theta_2|$. ■

Prueba del Lema 3.2. De forma similar al lema 3.1

$$\begin{aligned} v_t &= \theta_1 v_{t-1} 1(\varepsilon_{t-1} > r, \widehat{\varepsilon}_{t-1} > r) + \theta_2 v_{t-1} 1(\varepsilon_{t-1} < r, \widehat{\varepsilon}_{t-1} < r) + \\ &\quad (\theta_2 \widehat{\varepsilon}_{t-1} - \theta_1 \varepsilon_{t-1}) 1(\varepsilon_{t-1} > r, \widehat{\varepsilon}_{t-1} < r) + (\theta_1 \widehat{\varepsilon}_{t-1} - \theta_2 \varepsilon_{t-1}) 1(\varepsilon_{t-1} < r, \widehat{\varepsilon}_{t-1} > r). \end{aligned}$$

Usando que

$$\begin{aligned} |\theta_2 \widehat{\varepsilon}_{t-1} - \theta_1 \varepsilon_{t-1}| 1(\varepsilon_{t-1} > r, \widehat{\varepsilon}_{t-1} < r) &\leq [\max_i |\theta_i| |v_{t-1}| + |\theta_1 - \theta_2| r] 1(\varepsilon_{t-1} > r, \widehat{\varepsilon}_{t-1} < r) \\ |\theta_1 \widehat{\varepsilon}_{t-1} - \theta_2 \varepsilon_{t-1}| 1(\varepsilon_{t-1} < r, \widehat{\varepsilon}_{t-1} > r) &\leq [\max_i |\theta_i| |v_{t-1}| + |\theta_1 - \theta_2| r] 1(\varepsilon_{t-1} < r, \widehat{\varepsilon}_{t-1} > r), \end{aligned}$$

y el supuesto ASG.0 se puede probar que,

$$\begin{aligned} |v_t| &\leq \max_i |\theta_i| |v_{t-1}| + r |\theta_1 - \theta_2| [1(r < \varepsilon_{t-1} < r - v_{t-1}) + 1(r - v_{t-1} < \varepsilon_{t-1} < r)] \\ E|v_t| &\leq \max_i |\theta_i| E|v_{t-1}| + r |\theta_1 - \theta_2| \frac{1}{2} f^* E|v_{t-1}| \leq \lambda_1^* E|v_{t-1}|. \end{aligned}$$

Por lo tanto, siguiendo los mismos pasos que en el lema 3.1, obtenemos que $\lim_{t \rightarrow \infty} E|v_t^2| = 0$, y el proceso será invertible. ■

3.9.1. Consistencia y Tasa de Convergencia.

Caso del Tamaño

Como en el Teorema 2.2, denotamos $\phi = (\phi_1, \phi_2, \phi_3, \phi_4, \phi_5) = (\theta_1, \theta_2, \rho, \mu, r)$. Igualmente, $\phi^0 = (\phi_1^0, \phi_2^0, \phi_3^0, \phi_4^0, \phi_5^0) = (\theta_1^0, \theta_2^0, \rho^0, \mu^0, r^0)$ denotará el verdadero valor de los parámetros del proceso. Diremos que $\phi \in \Theta$, si $(\theta_1, \theta_2, r) \in \Theta_\theta$, $\rho \in (-1, 1)$ y $\mu \in [-\bar{\mu}, \bar{\mu}]$, donde Θ_θ está definido en el supuesto ASZ.1. El estimador CLS de ϕ^0 , $\hat{\phi}_T$, vendrá dado por:

$$\hat{\phi}_T = \arg \min_{\phi \in \Theta} Q_T(\phi).$$

Como en el caso Observable, tenemos la siguiente desigualdad,

$$\inf_{\phi \in \Theta | \rho(\phi, \phi^0) > \epsilon} Q_T(\phi) - Q_T(\phi^0) > \inf_{\phi \in \Theta | \rho(\phi, \phi^0) > \epsilon} \sum_{t=1}^T v_t^2(\phi, \phi^0) - \sup_{\phi \in \Theta | \rho(\phi, \phi^0) > \epsilon} \left| \sum_{t=1}^T 2\varepsilon_t v_t(\phi, \phi^0) \right|,$$

donde $v_t(\phi, \phi^0) = e_t(\phi) - \varepsilon_t$, $e_t(\phi) = x_t - \rho x_{t-1} - \mu + \theta_{t-1} e_{t-1}(\phi)$, $\theta_{t-1} = \theta_1 1(|e_{t-1}(\phi)| > r) + \theta_2 1(|e_{t-1}(\phi)| < r)$, $\varrho(\phi^*, \phi) = \max_i \{|\phi_i^* - \phi_i|\}$ y $\varepsilon_0 = 0$. Análogamente, la consistencia se deriva de las dos proposiciones, 2.15 y 2.16 del Anexo 2 en el Capítulo anterior, con las nuevas definiciones de Θ y $v_t(\phi, \phi^0)$.

La estrategia de la prueba será también la misma. En primer lugar definimos los procesos necesarios para acotar las variaciones producidas en $v_t(\phi, \phi^0)$ ante variaciones en ϕ , así como sus propiedades de momentos y dependencia.

Empezamos generalizando la definición de $A_{i,t}$ de la ecuación 3.3 de la siguiente manera,

$$\begin{aligned} A_{1,t-1}(\phi, \phi^*) &= \{|e_{t-1}(\phi)| > r, |e_{t-1}(\phi^*)| > r^*\} \\ A_{2,t-1}(\phi, \phi^*) &= \{|e_{t-1}(\phi)| < r, |e_{t-1}(\phi^*)| < r^*\} \\ A_{3,t-1}(\phi, \phi^*) &= \{|e_{t-1}(\phi)| > r, |e_{t-1}(\phi^*)| < r^*\} \\ A_{4,t-1}(\phi, \phi^*) &= \{|e_{t-1}(\phi)| < r, |e_{t-1}(\phi^*)| > r^*\} \end{aligned} \tag{3.6}$$

Además, definimos los siguientes procesos, $\bar{e}_t, v_t(\phi, \phi^*) = (e_t(\phi) - e_t(\phi^*))$ y $\bar{v}_t(\phi^*, \nu)$,

$$\bar{e}_t = |x_t| + |x_{t-1}| + \bar{\mu} + \bar{r} + (1 - \partial) \bar{e}_{t-1} \quad (3.7)$$

$$\begin{aligned} v_t(\phi, \phi^*) &= (\rho^* - \rho) x_{t-1} + (\mu^* - \mu) + \sum_{i=1}^2 (\theta_i - \theta_i^*) 1(A_{i,t-1}(\phi, \phi^*)) e_{t-1}(\phi^*) + \\ &+ [\theta_1 1(A_{1,t-1}(\phi, \phi^*)) + \theta_1^* 1(A_{4,t-1}(\phi, \phi^*))] v_{t-1}(\phi, \phi^*) + \\ &+ [\theta_2 1(A_{2,t-1}(\phi, \phi^*)) + \theta_1 1(A_{3,t-1}(\phi, \phi^*))] v_{t-1}(\phi, \phi^*) + \\ &+ (\theta_1 - \theta_2^*) 1(A_{3,t-1}(\phi, \phi^*)) e_{t-1}(\phi^*) + (\theta_2 - \theta_1^*) 1(A_{4,t-1}(\phi, \phi^*)) e_{t-1}(\phi) \end{aligned} \quad (3.8)$$

$$\begin{aligned} \bar{v}_t(\phi^*, \nu) &= \nu [|x_{t-1}| + 1 + (r^* + \nu) + 3\bar{e}_{t-1}] + \partial^* r^* [1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu))] + \\ &+ |\theta_{t-1}^*| \bar{v}_{t-1}(\phi^*, \nu) \end{aligned} \quad (3.9)$$

con $\partial^* = |\theta_2^* - \theta_1^*|$ y

$$\begin{aligned} &1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu)) = \\ &1[r^* - \nu - v_{t-1}(\phi^*, \phi^0) - \bar{v}_{t-1}(\phi^*, \nu) < \varepsilon_{t-1} < r^* + \nu - v_{t-1}(\phi^*, \phi^0) + \bar{v}_{t-1}(\phi^*, \nu)] + \\ &+ 1[-r^* - \nu - v_{t-1}(\phi^*, \phi^0) - \bar{v}_{t-1}(\phi^*, \nu) < \varepsilon_{t-1} < -r^* + \nu - v_{t-1}(\phi^*, \phi^0) + \bar{v}_{t-1}(\phi^*, \nu)]. \end{aligned} \quad (3.10)$$

Como es fácilmente comprobable,

$$\begin{aligned} \bar{e}_t &\geq \sup_{\phi} e_t(\phi) \\ 1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu)) &\geq \sup_{\phi \in \Theta | \rho(\phi, \phi^*) < \nu} |1(A_{3,t-1}(\phi, \phi^*)) + 1(A_{4,t-1}(\phi, \phi^*))| \\ \bar{v}_t(\phi^*, \nu) &\geq \sup_{\phi \in \Theta | \rho(\phi, \phi^*) < \nu} v_t(\phi, \phi^*) \end{aligned}$$

con \bar{e}_t no dependiendo de ϕ , y $\bar{A}_{i,t-1}(\phi^*, \nu)$ y $\bar{v}_t(\phi^*, \nu)$ dependiendo solo de ϕ^* y ν . Mostramos a continuación los resultados de momentos y dependencia de estos procesos.

Lema 3.3 *Bajo los supuestos G.0 y ASZ.1 se cumple que*

$$(a) \quad E(\bar{e}_t^{2\gamma}) \leq K_e < \infty$$

- (b) $E(\theta_t^{2\gamma} | F_{t-1}) \leq \lambda_2(\phi) < 1$
- (c) $E(\bar{v}_t^{2\gamma}(\phi^*, \nu)) \leq \nu K_v < \infty$
- (d) $E[1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu))] \leq \nu K_A < \infty$.

Prueba del Lema 3.3. Para el apartado (a), de la definición (3.7),

$$\|\bar{e}_t\|_{2\gamma} = \|x_t\|_{2\gamma} + \|x_{t-1}\|_{2\gamma} + \bar{\mu} + \bar{r} + (1 - \delta) \|\bar{e}_{t-1}\|_{2\gamma} \leq \frac{2\|x_t\|_{2\gamma} + \bar{\mu} + \bar{r}}{\delta} < \infty.$$

Para el apartado (b), teniendo en cuenta el supuesto ASZ.1,

$$\begin{aligned} E(\theta_t^{2\gamma} | \mathfrak{F}_{t-1}) &\leq E([(1 - \delta)^{2\gamma} 1(|e_t| > r) + 1(|e_t| < r)] | \mathfrak{F}_{t-1}) \leq \\ &\leq (1 - \delta)^{2\gamma} \left[1 - \sup_k E(1(|\varepsilon_t + k| < r)) \right] + \sup_k E(1(|\varepsilon_t + k| < r)) \\ &\leq \lambda_2(\phi) < 1. \end{aligned}$$

Para el apartado (c), usaremos el siguiente resultado,

$$E[1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu))] \leq 2f^*[\nu + E(\bar{v}_{t-1}(\phi^*, \nu))], \quad (3.11)$$

que se obtiene usando la LIE en la ecuación (3.10) y que $\bar{v}_{t-1}(\phi^*, \nu)$ es \mathfrak{F}_{t-1} -medible.

Ahora de la definición de $\bar{v}_t(\phi^*, \nu)$ en la ecuación (3.9), y los resultados anteriores,

$$\begin{aligned} E|\bar{v}_t(\phi^*, \nu)| &\leq \nu[E|x_{t-1}| + 1 + (r^* + \nu) + 3E(\bar{e}_{t-1})] \\ &\quad + 2\partial^* r^* f^*[\nu + E|\bar{v}_{t-1}(\phi^*, \nu)|] + \lambda_2(\phi^*) E|\bar{v}_{t-1}(\phi^*, \nu)| \\ &\leq \nu \frac{[E|x_{t-1}| + 1 + (r^* + \nu) + 3E(\bar{e}_{t-1}) + 2\partial^* r^* f^*]}{1 - \lambda_1} = \nu K_{v1}, \end{aligned}$$

usando en la última desigualdad que $\lambda_1(\phi^*) = \partial^* r^* 2f^* + \lambda_2(\phi^*) < \lambda_1 < 1$, por el supuesto

ASZ.1. Por lo tanto,

$$\begin{aligned} \|\bar{v}_t(\phi^*, \nu)\|_{2\gamma} &\leq \nu[\|x_{t-1}\| + 1 + (r^* + \nu) + 3\bar{e}_{t-1}\|_{2\gamma} + \lambda_2(\phi^*) \|\bar{v}_{t-1}(\phi^*, \nu)\|_{2\gamma} \\ &\quad + 2\partial^* r^* f^*[\nu + E|\bar{v}_{t-1}(\phi^*, \nu)|]^{1/2\gamma} \\ &\leq \nu^{1/2\gamma} K_v < \infty. \end{aligned}$$

Por último, la desigualdad (d), se obtiene directamente de la ecuación (3.11) y el apartado (c). ■

Lema 3.4 *Bajo los supuestos G.0 y ASZ.1,*

- (a) $e_t(\phi)$ es $L_{2\gamma}$ - acotado y L_2 - NED de tamaño $-\infty$ en ε_t , con constante $d_{e,t} = \varrho^{1/2}(\phi, \phi^0) K_e$
- (b) $\theta_t(e_t)$ es L_k - NED de tamaño $-\infty$ en ε_t , con constante $d_{\theta,t} = \varrho^{1/k}(\phi, \phi^0) K_\theta$ para $k \leq 4$.
- (c) \bar{e}_t es $L_{2\gamma}$ - acotado y L_2 - NED de tamaño $-\infty$ en ε_t , con constante $d_{\bar{e},t}$ finita, con K_e y K_θ finitas.

Prueba del Lema 3.4. La primera parte de los apartados (a) y (c) se obtiene del lema 3.3.

Para la prueba de L_2 - NED, empezaremos probando que x_t es L_2 - NED de tamaño $-\infty$ en ε_t ¹. Para ello, definimos

$$x_t^m = \sum_{j=0}^{m-1} (\rho^0)^j (\varepsilon_{t-j} - \theta_{t-j-1} \varepsilon_{t-j-1}) - \sum_{j=0}^{m-1} (\rho^0)^j \mu$$

y como en el caso del lema 2.4,

$$\|x_t - E(x_t / \mathfrak{F}_{t-m}^{t+m})\|_2 \leq \|x_t - x_t^m\|_2 \leq \|(\rho^0)^m x_{t-m}\|_2 \leq (\rho^0)^m \|x_{t-m}\|_2.$$

Tomando $v_{x,m} = (\rho^0)^m$ y $d_{x,t} = \|x_t\|_2$ queda probado. Para e_t , definimos,

$$e_t^m = \varepsilon_t - \theta_{t-1}^0 \varepsilon_{t-1} + (\rho^0 - \rho) x_{t-1}^{m-1} + (\mu^0 - \mu) + \theta_{t-1} (e_{t-1}^{m-1}) e_{t-1}^{m-1},$$

¹Si x_t es un proceso MA puro, entonces será ϕ - mixing, dado que ε_t es iid, y esta prueba no será necesaria.

donde $\theta_{t-1}(e_{t-1}^{m-1}) = \theta_1 1(|e_{t-1}^{m-1}| > r) + \theta_2 1(|e_{t-1}^{m-1}| < r)$. Tomando además $e_{t-m+1}^1 = \varepsilon_{t-m+1}$, e_t^m será \mathfrak{S}_{t-m}^{t+m} - medible. Así,

$$\begin{aligned} \|e_t - e_t^m\|_1 &\leq \|(\rho^0 - \rho)(x_{t-1} - x_{t-1}^{m-1})\|_1 + \| |\theta_{t-1}(e_{t-1})| (e_{t-1} - e_{t-1}^{m-1}) \|_1 + \\ &\quad + \partial r \|1(|e_{t-1}| > r, |e_{t-1}^{m-1}| < r) + 1(|e_{t-1}| < r, |e_{t-1}^{m-1}| > r)\|_1, \end{aligned}$$

y usando que $(e_{t-1} - e_{t-1}^{m-1})$ es \mathfrak{S}_{t-2} - medible, se tiene que

$$\|e_t - e_t^m\|_1 \leq (\rho^0 - \rho) \rho^{m-1} \|x_t\|_1 + \lambda_2(\phi) \|e_{t-1} - e_{t-1}^{m-1}\|_1 + 2\partial r f^* \|e_{t-1} - e_{t-1}^{m-1}\|_1.$$

Ahora, por el supuesto ASZ.1,

$$\|e_t - e_t^m\|_1 \leq m \max(\rho, \lambda_1(\phi))^{m-1} \max((\rho^0 - \rho) \|x_t\|_1, \|v_{t-m-1}(\phi, \phi^0)\|_1).$$

Sin embargo, para probar que el proceso es L_2 -NED necesitamos el segundo momento, para el que se cumple que

$$\begin{aligned} \|e_t - e_t^m\|_2 &\leq \|(\rho^0 - \rho)(x_{t-1} - x_{t-1}^{m-1})\|_2 + \| |\theta_{t-1}(e_{t-1})| (e_{t-1} - e_{t-1}^{m-1}) \|_2 + \\ &\quad \partial r \|1(|e_{t-1}| > r, |e_{t-1}^{m-1}| < r) + 1(|e_{t-1}| < r, |e_{t-1}^{m-1}| > r)\|_2 \\ &\leq (\rho^0 - \rho) \rho^{m-1} \|x_t\|_2 + \lambda_2(\phi) \|e_{t-1} - e_{t-1}^{m-1}\|_2 + \\ &\quad + \partial r (2f^* m \lambda_\rho^{m-1} \max((\rho^0 - \rho) \|x_t\|_1, \|v_{t-m-1}(\phi, \phi^0)\|_1))^{1/2} \\ &\leq m^{3/2} \lambda_\rho^{\frac{m-1}{2}} \left[\max((d_{1,e,t})^{1/2}, \|v_{t-m-1}(\phi, \phi^0)\|_2, (\rho^0 - \rho) \|x_t\|_2) K \right], \end{aligned}$$

con $\lambda_\rho = \max(\rho, \lambda_1) < 1$ y $d_{1,e,t} = \max((\rho^0 - \rho) \|x_t\|_1, \|v_{t-m-1}(\phi, \phi^0)\|_1)$. Usando el lema 3.3 y que $\|e_t - E(e_t / \mathfrak{S}_{t-m}^{t+m})\|_2 \leq \|e_t - e_t^m\|_2$ queda probado el apartado (a) para $v_{e,m} = m^{3/2} \lambda_\rho^{\frac{m-1}{2}}$ y

$$d_{e,t} = \left[\max((d_{1,e,t})^{1/2}, \|v_{t-m-1}(\phi, \phi^0)\|_2, (\rho^0 - \rho) \|x_t\|_2) K \right] \leq \varrho^{1/2}(\phi, \phi^0) K_e.$$

Para $\theta_t(e_t)$ y $k \leq 4$,

$$\begin{aligned} \|\theta_t(e_t) - \theta_t(e_t^m)\|_k &\leq \partial \|E1(|e_{t-1}| > r, |e_{t-1}^{m-1}| < r) + 1(|e_{t-1}| < r, |e_{t-1}^{m-1}| > r)\|_k \leq \\ &\leq \partial (2f^* m \lambda_\rho^{m-1} d_{1,e,t})^{1/k}, \end{aligned}$$

y tomado $v_m = m^{1/n} \lambda_\rho^{\frac{m-1}{k}}(\phi)$ y $d_{\theta,t} = [2\partial f^* d_{1,e,t}]^{1/n}$, completamos la prueba del apartado (b).

Finalmente, la prueba para \bar{e}_t se obtiene de forma similar teniendo en cuenta que x_t es $L_2 - NED$ en ε_t de tamaño $-\infty$ y constante finita. ■

Lema 3.5 *Bajo los supuestos G.0 y ASZ.1, $\bar{v}_t(\phi^*, \nu)$ es una $L_2 - mixingala$ en ε_t de tamaño $-1/2$ y constante $\max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2})\bar{K}_v$ con $\bar{K}_v < \infty$. Por lo tanto,*

$$E \left[\max_{1 \leq j \leq T} \left(\sum_{t=1}^j \chi_t \right)^2 \right] \leq K \sum_{t=1}^T \max(\varrho(\phi^*, \phi^0), \nu) \bar{K}_v,$$

para $\chi_t = |\varepsilon_t| \bar{v}_t(\phi^*, \nu) - E(|\varepsilon_t| \bar{v}_t(\phi^*, \nu))$.

Prueba del Lema 3.5. Para probar que $\bar{v}_t(\phi^*, \nu)$ es una $L_2 - mixingala$, primero probaremos que es un proceso $L_2 - NED$ en ε_t con constante $\max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2})\bar{K}_v$ de tamaño ∞ . Para ello, definimos la siguiente variable $\mathfrak{S}_{t-m}^{t+m} - medible$:

$$\begin{aligned} \bar{v}_t^m(\phi^*, \nu) &= \nu \left[|x_{t-1}^{m-1}| + 1 + (r^* + \nu) + 3\bar{e}_{t-1}^{m-1}(\phi) \right] + \theta_{t-1}^* (e_{t-1}^{m-1}) \bar{v}_{t-1}^{m-1}(\phi^*, \nu) \\ &\quad + \partial^* r^* \left[1 \left(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu) \right) \right], \end{aligned}$$

donde,

$$\begin{aligned} &1 \left(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu) \right) = \\ &= 1 \left[r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) - \nu - \bar{v}_{t-1}^{m-1}(\phi^*, \nu) < \varepsilon_{t-1} < r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) + \nu + \bar{v}_{t-1}^{m-1}(\phi^*, \nu) \right] + \\ &+ 1 \left[-r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) - \nu - \bar{v}_{t-1}^{m-1}(\phi^*, \nu) < \varepsilon_{t-1} < -r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) + \nu + \bar{v}_{t-1}^{m-1}(\phi^*, \nu) \right]. \end{aligned}$$

Usando que $|x_{t-1}|$ y $e_t(\phi)$ son secuencias $L_2 - NED$ en ε_t de tamaño $-\infty$, se obtiene que $v_t(\phi^*, \phi^0)$ también lo es, con constante $d_{e,t}$. Para ello, definimos

$$v_t^m(\phi^*, \phi^0) = e_t^m(\phi^*) - \varepsilon_t,$$

entonces,

$$\|v_t(\phi^*, \phi^0) - v_t^m(\phi^*, \phi^0)\|_2 \leq \|e_t - \varepsilon_t - e_t^m + \varepsilon_t\|_2 \leq m\lambda_\rho^{m-1} d_{e,t}.$$

Por lo tanto, de forma similar al lema 3.4,

$$\begin{aligned} \|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_1 &\leq \nu [\|x_{t-1} - x_{t-1}^{m-1}\|_1 + 3\|\bar{e}_{t-1}(\phi) - \bar{e}_{t-1}^{m-1}(\phi)\|_1] + \\ &+ \|\theta_{t-1}^*(e_{t-1}) - \theta_{t-1}^*(e_{t-1}^{m-1})\|_2 \|\bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_2 + \|\theta_{t-1}^*(e_{t-1}) [\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)]\|_1 + \\ &+ \partial^* r^* \left\| 1(\bar{A}_{3,t-1}(\phi^*, \nu)) - 1(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu)) - 1(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu)) \right\|_1. \end{aligned}$$

Además,

$$\begin{aligned} \left\| \sum_{i=3}^4 1(\bar{A}_{i,t-1}(\phi^*, \nu)) - 1(\bar{A}_{i,t-1}^{m-1}(\phi^*, \nu)) \right\|_1 &\leq 2f^* \|\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_1 + \\ &+ 2f^* \|v_{t-1}(\phi^*, \phi^0) - v_{t-1}^{m-1}(\phi^*, \phi^0)\|_1, \end{aligned}$$

y usando los resultados del lema 3.4,

$$\begin{aligned} \|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_1 &\leq \nu [\rho^{m-1} d_{x,t} + 3m\lambda_\rho^{m-1} d_{\bar{e},t}] + m^{1/2} \lambda_\rho^{\frac{m-1}{2}} d_{e,t} \|\bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_2 + \\ &+ (\lambda_2(\phi^*) + 2f^* \partial^* r^*) \|\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_1 + \\ &+ m\lambda_\rho^{m-1} d_{1,e,t} f^* \partial^* r^* \\ &\leq m^2 \lambda_\rho^{\frac{m-1}{2}} \max(\rho(\phi^*, \phi^0), \nu) K_v. \end{aligned}$$

El resto de la prueba sigue los mismos pasos que en el caso de $e_t(\phi)$, teniendo en cuenta que

$$\|\theta_{t-1}^*(e_{t-1}) - \theta_{t-1}^*(e_{t-1}^{m-1})\|_2 \|\bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_2 \leq \varrho^{1/4}(\phi^*, \phi^0) \nu^{1/4} K,$$

y por lo tanto,

$$\|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_2 \leq m^{\frac{5}{2}} \lambda_\rho^{\frac{m-1}{4}} \max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v.$$

con lo que $\bar{v}_t(\phi^*, \nu)$ es una secuencia $L_2 - NED$ de tamaño $-\infty$ en ε_t (que es ϕ -mixing) y constante $\max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v$. Con este resultado, y por el Teorema 2.10 del Anexo 1 en el Capítulo 2, queda probado el lema. ■

Con los resultados de estos tres lemas, ya podemos probar la consistencia de los estimadores, que se deriva de las proposiciones 2.15 y 2.16.

Prueba de la Proposición 2.15. Debemos probar que para todo $\eta > 0$ y $\epsilon > 0$

$$P \left[\sup_{\{\phi \in \Theta : \varrho(\phi, \phi^0) \geq \epsilon\}} \frac{1}{T} \left| \sum_{t=1}^T \varepsilon_t v_t(\phi, \phi^0) \right| > \eta \right] \xrightarrow{T \rightarrow \infty} 0.$$

Como en casos anteriores, usaremos $B_{\epsilon, j_l}^l(\phi_l) = \{\phi_l \in B_\epsilon / |\phi_l^{j_l} - \phi_l| \leq \nu\}$ para definir la partición de $B_\epsilon = \{\phi \in \Theta : \varrho(\phi, \phi^0) > \epsilon\}$, con $\phi_l^{j_l} = \phi_l^0 \pm \nu j_l$ para $l = 1, 2, \dots, 5$. Definiendo, $\ell_j = (j_1, j_2, j_3, j_4, j_5)$, para $j = 1, \dots, J$, obtenemos de nuevo que $B_\epsilon \subset \cup_{\ell_j=1}^J B_{\epsilon, \ell_j}(\phi)$, con $B_{\epsilon, \ell_j}(\phi) = \prod_{l=1}^5 B_{\epsilon, j_l}^l(\phi_l)$ y $\sup_{\phi \in B_{\epsilon, \ell_j}(\phi)} \varrho(\phi^{\ell_j}, \phi) = \nu$, para $\phi^{\ell_j} = (\phi_1^{j_1}, \phi_2^{j_2}, \phi_3^{j_3}, \phi_4^{j_4}, \phi_5^{j_5})$.

Acotamos el supremo con esta partición, para obtener que,

$$\sup_{\{\phi \in B_\epsilon\}} \frac{1}{T} \left| \sum_{t=1}^T \varepsilon_t v_t(\phi, \phi^0) \right| \leq \max_{\ell_j} \frac{1}{T} \left| \sum_{t=1}^T \varepsilon_t v_t(\phi^{\ell_j}, \phi^0) \right| + \max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}} \frac{1}{T} \left| \sum_{t=1}^T \varepsilon_t v_t(\phi^{\ell_j}, \phi) \right|. \quad (3.12)$$

Para el primer máximo de la ecuación (3.12), usamos que $\varepsilon_t v_t(\phi^j, \phi^0)$ es una *mds* y el lema 3.3, con lo que

$$P \left[\max_{\ell_j} \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_t v_t(\phi^{\ell_j}, \phi^0) \right| > \eta \right] \leq \frac{K_v \sigma_\varepsilon^2}{T \eta^2 \nu^4} \xrightarrow{T \rightarrow \infty} 0 \quad (3.13)$$

teniendo en cuenta que el $\max_{\ell_j} E(\varepsilon_t v_t(\phi^{\ell_j}, \phi^0))$ depende del índice máximo en ℓ_j , y siguiendo los mismos pasos que en la proposición 2.15 en el caso observable.

Para el segundo máximo de (3.12), usamos que $\sup_{\phi \in B_{\epsilon, j}} v_t(\phi^{\ell_j}, \phi) \leq \bar{v}_t(\phi^{\ell_j}, \nu)$ y el lema 3.5, para obtener

$$P \left[\max_{\ell_j} \frac{1}{T} \left| \sum_{t=1}^T [|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu) - E(|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu))] \right| > \frac{\eta}{2} \right] \leq \frac{K_v \sigma_\varepsilon^2}{T \eta^2 \nu^3} \xrightarrow{T \rightarrow \infty} 0 \quad (3.14)$$

$$\max_{\ell_j} E(|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu)) \leq \nu K_{v,1} \quad (3.15)$$

De las desigualdades (3.14) y (3.15) para ν suficientemente pequeño, es directo probar que

$$P \left[\max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}} \left| \frac{1}{T} \sum_{t=1}^T \varepsilon_t v_t(\phi^{\ell_j}, \nu) \right| > \eta \right] \leq \frac{K(\nu)}{T \eta^2} \xrightarrow{T \rightarrow \infty} 0 \quad (3.16)$$

Con (3.13) y (3.16) queda probada la proposición 2.15. ■

Prueba de la Proposición 2.16. Como en la proposición 2.16 en el caso Observable, tenemos,

$$W_{k(\iota)}(\phi, \phi^0) = v_{k(\iota)-1}^2(\phi, \phi^0) + v_{k(\iota)}^2(\phi, \phi^0) \quad \text{con } k(\iota) = 2\iota,$$

y definiendo $\delta_{\phi, l}^0 = \phi_l - \phi_l^0$, podemos escribir $v_{k(\iota)}(\phi, \phi^0)$ como,

$$v_{k(\iota)}(\phi, \phi^0) = \delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1} + \theta_{k(\iota)-1} v_{k(\iota)-1}(\phi, \phi^0),$$

con $\theta_{k(\iota)-1} = \theta_{k(\iota)-1}(e_{k(\iota)-1}(\phi))$ y $\theta_{k(\iota)-1}^0 = \theta_{k(\iota)-1}^0(\varepsilon_{k(\iota)-1})$. Condicionando en los sucesos,

$$|v_{k(\iota)-1}(\phi, \phi^0)| \leq \frac{1}{2} \min \{ |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1}|, \bar{\epsilon}/n \}$$

para $4 \leq n < \infty$ y $\bar{\epsilon} = \min(\bar{r} - r^0, r^0 - \underline{r}, \epsilon)$, y en los sucesos,

$$\begin{aligned} |\delta_{\phi, 5}^0| &< \bar{\epsilon}/2 \\ \max_{l \leq 4} |\delta_{\phi, l}^0| &\geq \epsilon^* > 0, \end{aligned}$$

con $\epsilon^* > 0$, que definiremos más adelante, obtenemos que

$$\begin{aligned} W_{k(\iota)}(\phi, \phi^0) &\geq \min \left(\frac{\bar{\epsilon}^2}{n^2}, \vartheta_{k(\iota), 1}^2(\delta_{\phi}^0) \right) \\ \vartheta_{k(\iota), 1}^2(\delta_{\phi}^0) &= \frac{1}{2} 1(|\varepsilon_{k(\iota)-1}| > \bar{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 1}^0 \varepsilon_{k(\iota)-1}|^2 + \\ &\quad \frac{1}{2} 1(|\varepsilon_{k(\iota)-1}| < \underline{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 2}^0 \varepsilon_{k(\iota)-1}|^2. \end{aligned}$$

Dado el supuesto ASZ.2, $E(M_{k(\iota)-1}(\bar{r}))$ y $E(M_{k(\iota)-1}(\underline{r}))$ son definidas positivas, y,

$$E[\vartheta_{k(\iota), 1}^2(\delta_{\phi}^0)] \geq \max_{i \leq 4} |\delta_{\phi, i}^0|^2 a_1 \geq \epsilon^* a_1,$$

con $a_1 > 0$. Condicionando en los sucesos complementarios respecto a $\delta_{\phi, l}^0$,

$$\begin{aligned} \delta_{\phi, 5}^0 &> \bar{\epsilon}/2 \\ \max_{l \leq 4} |\delta_{\phi, l}^0| &< \epsilon^* > 0, \end{aligned}$$

el caso, $\delta_{\phi,5}^0 < -\bar{\epsilon}/2$ es similar, se obtiene que,

$$W_{k(\iota)}(\phi, \phi^0) \geq \min \left(\frac{\bar{\epsilon}}{n^2}, \vartheta_{k(\iota),2}^2(\delta_\phi^0) \right)$$

$$\vartheta_{k(\iota),2}^2(\delta_\phi^0) = \frac{1}{2} 1 \left(r^0 < \varepsilon_{k(\iota)-1} < r^0 + (\delta_{\phi,5}^0 - \bar{\epsilon}/4) \right) \left| \epsilon^* x_{k(\iota)-1} + \epsilon^* + (\delta^0 - \epsilon^*) \varepsilon_{k(\iota)-1} \right|^2.$$

Para ϵ^* suficientemente pequeño, se tiene que,

$$E(\vartheta_{k(\iota),2}^2(\delta_\phi^0)) \geq \frac{1}{4} \delta_{\phi,5}^0 f_* K_\sigma \geq \frac{1}{4} \bar{\epsilon} f_* K_\sigma > 0.$$

Por lo tanto,

$$\inf_{\{\phi \in \Theta: \rho(\phi, \phi^0) \geq \epsilon\}} E(W_{k(\iota)}(\phi, \phi^0)) > \min_{j=1,2} E \left(\min \left(\frac{\bar{\epsilon}}{n^2}, \vartheta_{k(\iota),j}^2(\delta_\phi^0) \right) \right) = \min(\bar{\epsilon}, \epsilon^*, \epsilon) K_w > 0.$$

Esta última desigualdad se obtiene para n suficientemente grande, tal que

$$\min_{j=1,2} E(\vartheta_{k(\iota),j}^2) \geq \frac{\bar{\epsilon}}{n^2}, \text{ dado que } E(\vartheta_{k(\iota),j}^4) < \infty.$$

Ahora, la prueba de la proposición se seguirá de una ULLN para $T^{-1} \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0)$.

Para ello acotamos las variaciones respecto a ϕ de $W_{k(\iota)}(\phi, \phi^0)$ y obtenemos las propiedades de dependencia y momentos de las cotas. Así,

$$\begin{aligned} \|W_{k(\iota)}(\phi^{\ell_j}, \phi^0)\|_2 &\leq \|v_{k(\iota)-1}^2(\phi^{\ell_j}, \phi^0)\|_2 + \|v_{k(\iota)}^2(\phi^{\ell_j}, \phi^0)\|_2 \\ &\leq \rho^{1/2}(\phi^{\ell_j}, \phi^0) K_{w,1} \\ \sup_{\phi \in B_{\epsilon, \ell_j}} |W_{k(\iota)}(\phi^{\ell_j}, \phi^0) - W_{k(\iota)}(\phi, \phi^0)| &\leq \sum_{i=0}^1 \bar{v}_{k(\iota)-i}(\phi^{\ell_j}, \nu) [2v_{k(\iota)-i}(\phi^{\ell_j}, \phi^0) + \bar{v}_{k(\iota)-i}(\phi^{\ell_j}, \nu)] \\ E| \sup_{\phi \in B_{\epsilon, \ell_j}} W_{k(\iota)}(\phi^{\ell_j}, \phi^0) - W_{k(\iota)}(\phi, \phi^0) | &\leq \nu^{1/2} K_W. \end{aligned}$$

Usando los Teoremas 2.13 y 2.14 en el Anexo 1 del capítulo anterior se obtiene que $W_{k(\iota)}(\phi, \phi^0)$ y $\sum_{i=0}^1 \bar{v}_{k(\iota)-i}(\phi^{\ell_j}, \nu) [2v_{k(\iota)-i}(\phi^{\ell_j}, \phi^0) + \bar{v}_{k(\iota)-i}(\phi^{\ell_j}, \nu)]$ son L_2 -mixingales de tamaño $-1/2$ y constantes $\varrho^{1/2}(\phi, \phi^0) K_W$ y $\nu^{1/2} K_W$ respectivamente. Con estos resultados, y de forma similar a la proposición 2.19, obtenemos la ULLN, de forma que para $\forall \eta > 0$ y $\varepsilon > 0 \exists T_\epsilon$ tal que para $\forall T > T_\epsilon$,

$$P \left[\sup_{\{\phi \in \Theta: \rho(\phi, \phi^0) \geq \epsilon\}} T^{-1} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) - E(W_{k(\iota)}(\phi, \phi^0)) \right| > \eta \right] \leq \varepsilon$$

con lo que queda probada la proposición. ■

Prueba del Theorema 3.1.

Similar al caso observable, definimos,

$$B_{\Delta, \epsilon} = \left\{ \phi \in \Theta : \frac{\Delta}{T} \leq |\phi_i - \phi_i^0| = \delta_{\phi, i}^0 \leq \epsilon \text{ para } i = 1, \dots, 5 \right\},$$

que es acotado. La prueba se realizará para cualquier $\epsilon > 0$ ya que dada la consistencia de $\hat{\phi}_{i, T}$, para T_ϵ^* suficientemente grande, $\hat{\phi}_{i, T} \in B_{\Delta, \epsilon}$ para todo $T \geq T_\epsilon^*$ con probabilidad tan cercana a 1 como queramos. La prueba del teorema se obtendrá de las siguientes dos proposiciones.

Proposición 3.7 *Para todo $\eta, \epsilon > 0 \exists \Delta, T_\epsilon^* < \infty$ tal que para todo $T \geq T_\epsilon^*$*

$$P \left[\sup_{\{\phi \in B_{\Delta, \epsilon}\}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t(v_t(\phi, \phi^0))}{T\varrho(\phi, \phi^0)} \right| > \eta \right] \leq \epsilon.$$

Proposición 3.8 *Para todo $\epsilon > 0 \exists \Delta, T_\epsilon^* < \infty$ y $\eta > 0$ tal que para todo $T \geq T_\epsilon^*$*

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T\varrho(\phi, \phi^0)} > \eta \right] \geq 1 - \epsilon.$$

La principal diferencia con las proposiciones 2.17 y 2.18 es que, en este caso, los procesos son estandarizados por $T\varrho(\phi, \phi^0)$ y no por $T\bar{\varrho}^2(\phi, \phi^0)$, dado que todos los parámetros convergen a la misma tasa, T .

Prueba de la Proposición 3.7. De forma similar a la proposición 2.17, realizaremos la prueba para el caso $\phi_l > \phi_l^0$ y $l = 1, 2, \dots, 5$. El resto de casos, se obtiene siguiendo los

mismos pasos. Así, definimos la siguiente partición para cada subespacio paramétrico,

$$B_{\epsilon, j_l}^l(\Delta, \phi_l) = \left\{ \phi_l \in B_{\Delta, \epsilon}^+ \middle/ 0 \leq \frac{\phi_l^{j_l} - \phi_l}{\phi_l^{j_l} - \phi_l^0} \leq (b-1) \right\} \text{ con } b > 1, \phi_l^{j_l} = \phi_l^0 + \frac{\Delta}{T} b^{j_l} \text{ y } j_l =$$

$1, \dots, J_l$, para $l = 1, \dots, 5$. Por lo tanto, definiendo como en casos anteriores, $\ell_j = (j_1, j_2, j_3, j_4, j_5)$,

para $j = 1, \dots, J$ y $J = \prod_{l=1}^5 J_l$, es directo probar que $B_{\Delta, \epsilon} \subset \cup_{j=1}^J B_{\epsilon, \ell_j}^\Delta(\phi)$, con $B_{\epsilon, \ell_j}^\Delta(\phi) =$

$\prod_{l=1}^5 B_{\epsilon, j_l}^l(\Delta, \phi_l)$. Para este caso, y definiendo $\phi^{\ell_j} = (\phi_1^{j_1}, \phi_2^{j_2}, \phi_3^{j_3}, \phi_4^{j_4}, \phi_5^{j_5})$:

$$\sup_{\phi \in B_{\epsilon, \ell_j}^\Delta(\phi)} \varrho(\phi^{\ell_j}, \phi) = \nu_\Delta(\ell_j) = \varrho(\phi^{\ell_j}, \phi^0)(b-1).$$

Ahora, usando que la esperanza de $\varepsilon_t v_t(\phi^{\ell_j}, \phi^0)$ y $|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu_\Delta(\ell_j))$ dependen únicamente de $\varrho(\phi^{\ell_j}, \phi^0)$ y $\nu_\Delta(\ell_j)$ respectivamente, podemos distinguir cinco casos en el \max_{ℓ_j} , como en el caso Observable, dependiendo de que índice es máximo en ℓ_j . Por lo tanto, usando el resultado del lema 3.5, tenemos las siguientes tres desigualdades,

$$\begin{aligned}
 P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T \varepsilon_t (v_t(\phi^{\ell_j}, \phi^0))}{T \varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta \right] &\leq 5 \sum_{j=1}^{\max_l J_l} \frac{j^4 \left(\frac{\Delta}{T} b^j\right) \sigma_\varepsilon^2 K_v}{T \eta^2 \left(\frac{\Delta}{T} b^j\right)^2} \\
 &\leq \frac{5 \sigma_\varepsilon^2 K_v}{\Delta \eta^2} \sum_{j=1}^{\infty} \frac{j^4}{b^j} \\
 P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T |\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu_\Delta(\ell_j)) - E[|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu_\Delta(\ell_j))]}{T \varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta \right] &\leq \frac{5 \sigma_\varepsilon^2 K_v}{\Delta \eta^2} \sum_{j=1}^{\infty} \frac{j^4}{b^j} \\
 \max_{\ell_j} \frac{E[|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu_\Delta(\ell_j))]}{\varrho(\phi^{\ell_j}, \phi^0)} &\leq \sigma_\varepsilon K_v^{1/2} (b-1).
 \end{aligned}$$

En la primera desigualdad, usamos que se trata de la suma de una mds . En la segunda, usamos los resultados del lema 3.5. La tercera desigualdad, se obtiene de forma directa por el lema 3.3. Dado que $\max_{\ell_j} \frac{E[|\varepsilon_t| \bar{v}_t(\phi^{\ell_j}, \nu_\Delta(\ell_j))]}{\varrho(\phi^{\ell_j}, \phi^0)}$ es tan pequeño como queramos para b suficientemente cercano a 1,

$$P \left[\sup_{\{\phi \in B_{\Delta, \epsilon}\}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t (v_t(\phi, \phi^0))}{T \varrho(\phi, \phi^0)} \right| > 3\eta \right] \leq \frac{10 \sigma_\varepsilon^2 K_v}{\Delta \eta^2} \sum_{j=1}^{\infty} \frac{j^4}{b^j}. \quad (3.17)$$

Con este resultado queda probada la proposición 3.7, dado que $\sum_{j=1}^{\infty} \frac{j^4}{b^j} < \infty$, y por lo tanto para Δ suficientemente grande, la probabilidad de la ecuación (3.17) es tan pequeña como queramos. ■

Prueba de la Proposición 3.8. Debemos probar que, para todo $\varepsilon > 0 \exists \Delta, T_\epsilon^* < \infty$ y $\eta > 0$ tal que para todo $T \geq T_\epsilon^*$

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho(\phi, \phi^0)} > \eta \right] \geq 1 - \varepsilon,$$

o lo que es lo mismo,

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} 2 \frac{\sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho(\phi, \phi^0)} > 2\eta \right] \geq 1 - \varepsilon.$$

Como en la prueba de la proposición 2.16, definimos $W_{k(\iota)}(\phi, \phi^0) = (v_{k(\iota)}(\phi, \phi^0))^2 + (v_{k(\iota)-1}(\phi, \phi^0))^2$ para $k(\iota) = 2\iota$, de forma que

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{2 \sum_{t=1}^T (v_t(\phi, \phi^0))^2}{T \varrho(\phi, \phi^0)} > 2\eta \right] \geq P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) + W_{k(\iota)-1}(\phi, \phi^0)}{T \varrho(\phi, \phi^0)} > 2\eta \right].$$

Por consistencia, $\max_{l \leq 5} |\delta_{\phi, l}^0| \leq \epsilon$ para T suficientemente grande, y de la definición de $v_t(\phi, \phi^0)$, condicionando en los sucesos

$$E = \left\{ |v_{k(\iota)-1}(\phi, \phi^0)| > \frac{1}{2} |(\delta^0 - \epsilon)(r^0 - \epsilon)| [1(A_{3, k(\iota)-1}(\phi, \phi^0)) + 1(A_{4, k(\iota)-1}(\phi, \phi^0))] \right\},$$

y su complementario, E^c , tenemos el siguiente límite inferior para $W_{k(\iota)}(\phi, \phi^0)$,

$$W_{k(\iota)}(\phi, \phi^0) \geq \left[\frac{1}{2} (\delta^0 - \epsilon)(r^0 - \epsilon) \right]^2 [1(A_{3, k(\iota)-1}(\phi, \phi^0)) + 1(A_{4, k(\iota)-1}(\phi, \phi^0))] - \max_{l \leq 4} |\delta_{\phi, l}^0| (|x_{k(\iota)-1}| + 1 + |\varepsilon_{k(\iota)-1}|) |\bar{e}_{k(\iota)-1}| (1(A_{3, k(\iota)-1}(\phi, \phi^0)) + 1(A_{4, k(\iota)-1}(\phi, \phi^0))).$$

Por lo tanto, definiendo

$$\begin{aligned} \vartheta_{k(\iota), 2}^2(\delta_{\phi}^0) &= \sum_{j=1}^2 \left[\frac{1}{2} (\delta^0 - \epsilon)(r^0 - \epsilon) \right]^2 [1(A_{3, k(\iota)-j}(\phi, \phi^0)) + 1(A_{4, k(\iota)-j}(\phi, \phi^0))] \\ \vartheta_{k(\iota), 1}^2(\phi, \phi^0) &= \sum_{j=1}^2 (|x_{k(\iota)-j}| + 1 + |\varepsilon_{k(\iota)-j}|) |\bar{e}_{k(\iota)-j}| [1(A_{3, k(\iota)-j}(\phi, \phi^0)) + 1(A_{4, k(\iota)-j}(\phi, \phi^0))], \end{aligned}$$

se tiene que,

$$\begin{aligned} \inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} W_{k(\iota)}(\phi, \phi^0) + W_{k(\iota)-1}(\phi, \phi^0)}{T \varrho(\phi, \phi^0)} &\geq \inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 2}^2(\phi, \phi^0)}{T \varrho(\phi, \phi^0)} - \\ &\quad - \sup_{\{\phi \in B_{\Delta, \epsilon}\}} \max_{l \leq 4} |\delta_{\phi, l}^0| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 1}^2(\phi, \phi^0)}{T \varrho(\phi, \phi^0)}. \end{aligned}$$

Ahora, usando que $E \{ (|x_{t-1}| + 1 + |\varepsilon_{t-1}|) |\bar{e}_{t-1}| [1(A_{3, t-1}(\phi, \phi^0)) + 1(A_{4, t-1}(\phi, \phi^0))] \} \leq \varrho(\phi, \phi^0) K_{\vartheta}$ es fácil probar que para $\forall \eta > 0 \exists T_{\epsilon}^*$ suficientemente grande, tal que para todo $T > T_{\epsilon}^*$,

$$P \left[\sup_{\{\phi \in B_{\Delta, \epsilon}\}} \max_{l \leq 4} |\delta_{\phi, l}^0| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 1}^2(\phi, \phi^0)}{T \varrho(\phi, \phi^0)} > \eta \right] \leq \frac{\varepsilon}{2},$$

dada la consistencia de $\widehat{\phi}_T$. Así, para probar la proposición 3.8, bastará con probar

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 2}^2(\phi, \phi^0)}{T \varrho(\phi, \phi^0)} > \eta \right] \geq 1 - \frac{\varepsilon}{2},$$

o lo que es lo mismo,

$$P \left[\sup_{\{\phi \in B_{\Delta, \epsilon}\}} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 2}^2(\phi, \phi^0) - E \left(\vartheta_{k(\iota), 2}^2(\phi, \phi^0) \right)}{T \varrho(\phi, \phi^0)} \right| > \eta^* \right] \leq \frac{\varepsilon}{2},$$

si se cumple que $\inf_{\phi \in B_{\Delta, \epsilon}} \frac{E(\vartheta_{k(\iota), 2}^2(\phi, \phi^0))}{2\varrho(\phi, \phi^0)} - \eta^* > \eta$, para η^* suficientemente pequeño. Por lo tanto, en primer lugar, probamos que

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} E \left(\frac{\vartheta_{k(\iota), 2}^2(\phi, \phi^0)}{\varrho(\phi, \phi^0)} \right) > a > 0$$

para aplicar posteriormente la ULLN a $\frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 2}^2(\phi, \phi^0)}{T \varrho(\phi, \phi^0)}$. Usando la LIE,

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} E \left(\vartheta_{k(\iota), 2}^2(\phi, \phi^0) \right) \geq [(\delta^0 - \epsilon)(r^0 - \epsilon)] \left[f_* \left(2|\delta_{\phi, 5}^0| + \inf_{\{\phi \in B_{\Delta, \epsilon}\}} E \left(\Upsilon_{k(\iota)}(\phi, \phi^0, K) \right) \right) \right]$$

con $f_* = \min_{|e| \leq \bar{r}} f(e) > 0$, $K = \min(r^0 - \underline{r} - \epsilon, \bar{r} - r^0 - \epsilon)$ y

$$E \left(\Upsilon_{k(\iota)}(\phi, \phi^0, K) \right) = E \left[\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| \mathbf{1} \left(\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| < K \right) \right].$$

Si $|\delta_{\phi, 5}^0| = \varrho(\phi, \phi^0) > \frac{\Delta}{T}$, se obtiene el resultado deseado. Por lo tanto, nos centramos en el caso, $|\delta_{\phi, 5}^0| < \varrho(\phi, \phi^0)$. Condicionando a los sucesos

$$\begin{aligned} |v_{k(\iota)-2}(\phi, \phi^0)| &\geq \frac{1}{2} \mathbf{1}(|\varepsilon_{k(\iota)-1}| > \bar{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 1}^0 \varepsilon_{k(\iota)-1}| + \\ &\quad \frac{1}{2} \mathbf{1}(|\varepsilon_{k(\iota)-1}| < \underline{r}) |\delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + \delta_{\phi, 2}^0 \varepsilon_{k(\iota)-1}|, \end{aligned}$$

es fácil probar que,

$$\begin{aligned} & 1 \left(\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| < K \right) [|v_{k(\iota)-1}(\phi, \phi^0)| + |v_{k(\iota)-2}(\phi, \phi^0)|] \geq \\ & 1 \left(\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| < K \right) \frac{1}{2} 1(|\varepsilon_{k(\iota)-1}| > \bar{r}) |\delta_{\phi,3}^0 x_{k(\iota)-1} + \delta_{\phi,4}^0 + \delta_{\phi,1}^0 \varepsilon_{k(\iota)-1}| + \\ & + 1 \left(\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| < K \right) \frac{1}{2} 1(|\varepsilon_{k(\iota)-1}| < \underline{r}) |\delta_{\phi,3}^0 x_{k(\iota)-1} + \delta_{\phi,4}^0 + \delta_{\phi,2}^0 \varepsilon_{k(\iota)-1}|. \end{aligned}$$

Además, con el supuesto ASZ.2 y dado que

$$\left\| 1 \left(\sum_{i=1}^2 |\bar{v}_{k(\iota)-i}(\phi^0, \epsilon)| > K \right) \right\|_2 \leq \frac{\epsilon^{1/2} K_v^{1/2}}{K},$$

se tiene que para ϵ suficientemente pequeño,

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} E \left[\sum_{i=1}^2 \frac{|v_{k(\iota)-i}(\phi, \phi^0)|}{\varrho(\phi, \phi^0)} 1 \left(\sum_{i=1}^2 |v_{k(\iota)-i}(\phi, \phi^0)| < K \right) \right] > a > 0,$$

que garantiza

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} E \left(\frac{\vartheta_{k(\iota),2}^2(\phi, \phi^0)}{\varrho(\phi, \phi^0)} \right) > [(\delta^0 - \epsilon)(r^0 - \epsilon)] [f_* \min(2, a)] > 0.$$

A continuación probamos la ULLN. Como en casos anteriores, usamos $B_{\epsilon, \ell_j}^{\Delta}(\phi)$ para acotar el supremo, tal que,

$$\begin{aligned} & \sup_{\{\phi \in B_{\Delta, \epsilon}\}} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi, \phi^0) - E \left(\vartheta_{k(\iota),2}^2(\phi, \phi^0) \right)}{T \varrho(\phi, \phi^0)} \right| \leq \\ & \quad \max_{\ell_j} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) - E \left(\vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) \right)}{T \varrho(\phi^{\ell_j}, \phi^0)} \right| + \\ & \quad \max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) - E \left(\vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) \right)}{T \varrho(\phi^{\ell_j}, \phi^0)} - \frac{\vartheta_{k(\iota),2}^2(\phi, \phi^0) - E \left(\vartheta_{k(\iota),2}^2(\phi, \phi^0) \right)}{T \varrho(\phi, \phi^0)} \right| \end{aligned}$$

Definiendo,

$$\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)) = [(\delta^0 - \epsilon)(r^0 - \epsilon)] \sum_{j=1}^2 \sum_{i=3}^4 1(\bar{A}_{i,k(\iota)-j}(\phi^0, \nu_\Delta(\ell_j))),$$

tenemos,

$$\begin{aligned} & P \left[\sup_{\{\phi \in B_{\Delta,\epsilon}\}} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi, \phi^0) - E(\vartheta_{k(\iota),2}^2(\phi, \phi^0))}{T\varrho(\phi, \phi^0)} \right| > \eta \right] \leq \\ & P \left[\max_{\ell_j} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) - E(\vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0))}{T\varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta/2 \right] + \\ & + P \left[\max_{\ell_j} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)) - E(\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)))}{T\varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta/4 \right], \end{aligned}$$

usando además que $\|\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j))\|_2 \leq \nu_\Delta(\ell_j) K_W$, y por lo tanto, para b suficientemente cercano a 1

$$\sup_{\{\phi \in B_{\Delta,\epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} E(\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)))}{T\varrho(\phi^{\ell_j}, \phi^0)} \leq \frac{(b-1)K_W}{2} \leq \eta/4$$

Finalmente, dados los resultados de L_2 -mixingala para $\vartheta_{k(\iota),2}^2(\phi, \phi^0)$ y $\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j))$, que se obtienen de la forma directa a través de los resultados del lema 3.4, es fácil probar que,

$$\begin{aligned} & P \left[\max_{\ell_j} \left| \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0) - E(\vartheta_{k(\iota),2}^2(\phi^{\ell_j}, \phi^0))}{T\varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta/2 \right] \leq \sum_{j=1}^{\infty} \frac{j^4 K'_W}{\Delta b^j \eta} \\ & P \left[\max_{\ell_j} \left| \sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \frac{\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)) - E(\bar{\vartheta}_{k(\iota),2}^2(\phi^0, \nu_\Delta(\ell_j)))}{T\varrho(\phi^{\ell_j}, \phi^0)} \right| > \eta/4 \right] \leq \sum_{j=1}^{\infty} \frac{j^4 K'_W}{\Delta b^j \eta}. \end{aligned}$$

Ambas probabilidades pueden ser tan pequeñas como queramos para Δ suficientemente grande. Por lo tanto,

$$P \left[\inf_{\{\phi \in B_{\Delta,\epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota),2}^2(\phi, \phi^0)}{T\varrho(\phi, \phi^0)} > \inf_{\{\phi \in B_{\Delta,\epsilon}\}} \frac{E(\vartheta_{k(\iota),2}^2(\phi, \phi^0))}{2\varrho(\phi, \phi^0)} - \eta^* \right] > 1 - \varepsilon,$$

para Δ suficientemente grande, con

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{E \left(\vartheta_{k(\iota), 2}^2 (\phi, \phi^0) \right)}{2 \varrho (\phi, \phi^0)} - \eta^* \geq \frac{[(\delta^0 - \epsilon) (r^0 - \epsilon)]^2 f_* \min (2, a)}{2} - \eta^* > 0,$$

para $\eta^* > 0$ suficientemente pequeño. El resultado de la proposición se obtiene usando que

$$\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{t=1}^T (v_t (\phi, \phi^0))^2}{T \varrho (\phi, \phi^0)} > \inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 2}^2 (\phi, \phi^0)}{T \varrho (\phi, \phi^0)} - \sup_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{\iota=2}^{\lfloor \frac{T}{2} \rfloor} \vartheta_{k(\iota), 1}^2 (\phi, \phi^0)}{T \varrho (\phi, \phi^0)}.$$

Por lo tanto,

$$P \left[\inf_{\{\phi \in B_{\Delta, \epsilon}\}} \frac{\sum_{t=1}^T (v_t (\phi, \phi^0))^2}{T \varrho (\phi, \phi^0)} > \frac{[(\delta - \epsilon) (r^0 - \epsilon)]^2 f_* K_1}{2} - \eta \right] > 1 - \varepsilon.$$

■

Caso del Signo.

La prueba para este caso es paralela a la del caso del tamaño. La principal diferencia aparece en la identificabilidad y tasa de convergencia cuando $r^0 = 0$. Como en el caso anterior, empezamos por generalizar la definición de $A_{i,t}$ para el caso del signo. Así, definimos los sucesos $A_{i,t-1} (\phi, \phi^*)$ de la siguiente manera,

$$\begin{aligned} A_{1,t-1} (\phi, \phi^*) &= \{e_{t-1} (\phi) > r, e_{t-1} (\phi^*) > r^*\} & A_{2,t-1} (\phi, \phi^*) &= \{e_{t-1} (\phi) < r, e_{t-1} (\phi^*) < r^*\} \\ A_{3,t-1} (\phi, \phi^*) &= \{e_{t-1} (\phi) > r, e_{t-1} (\phi^*) < r^*\} & A_{4,t-1} (\phi, \phi^*) &= \{e_{t-1} (\phi) < r, e_{t-1} (\phi^*) > r^*\} \end{aligned} \quad (3.18)$$

donde $e_t (\phi) = x_t - \rho x_{t-1} - \mu + \theta_{t-1} e_{t-1} (\phi)$, $\theta_{t-1} = \theta_1 1 (e_{t-1} (\phi) > r) + \theta_2 1 (e_{t-1} (\phi) < r)$.

Como en el caso del tamaño, $v_t (\phi, \phi^*) = (e_t (\phi) - e_t (\phi^*))$. Si $|r^0| > 0$, definimos

$$\begin{aligned} \bar{v}_t (\phi^*, \nu) &= \nu [|x_{t-1}| + 1 + (r^* + \nu) + 3\bar{e}_{t-1}] + \partial^* r^* [1 (\bar{A}_{3,t-1} (\phi^*, \nu)) + 1 (\bar{A}_{4,t-1} (\phi^*, \nu))] + \\ &+ \max_i |\theta_i^*| \bar{v}_{t-1} (\phi^*, \nu), \end{aligned} \quad (3.19)$$

con $\partial^* = |\theta_2^* - \theta_1^*|$, $\bar{e}_t = |x_t| + |x_{t-1}| + \bar{\mu} + \lambda_2 \bar{e}_{t-1}$ y

$$\begin{aligned} & 1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu)) = \\ & 1[r^* - \nu - v_{t-1}(\phi^*, \phi^0) - \bar{v}_{t-1}(\phi^*, \nu) < \varepsilon_{t-1} < r^* - v_{t-1}(\phi^*, \phi^0)] + \\ & + 1[-r^* - v_{t-1}(\phi^*, \phi^0) < \varepsilon_{t-1} < -r^* + \nu - v_{t-1}(\phi^*, \phi^0) + \bar{v}_{t-1}(\phi^*, \nu)]. \end{aligned} \quad (3.20)$$

Si r^0 es conocido e igual a 0, nos centraremos en los parámetros $\beta = (\theta_1, \theta_2, \rho, \mu)$, definiendo $\bar{v}_t(\beta^*, \nu)$ como,

$$\bar{v}_t(\beta^*, \nu) = \nu[|x_{t-1}| + 1 + 3\bar{e}_{t-1}] + \max_i |\theta_i^*| \bar{v}_{t-1}(\beta^*, \nu). \quad (3.21)$$

Usando que

$$\begin{aligned} (\theta_i e_{t-1}(\phi) - \theta_j^* e_{t-1}(\phi^*)) 1(A_{k,t-1}(\phi, \phi^*)) & \leq \max_{l=i,j} |\theta_l^*| v_{t-1}(\phi, \phi^*) 1(A_{k,t-1}(\phi, \phi^*)) + \\ & + |\theta_i - \theta_j^*| \max(|r|, |r^*|) 1(A_{k,t-1}(\phi, \phi^*)), \end{aligned} \quad (3.22)$$

es fácil obtener que

$$\begin{aligned} \bar{e}_t & \geq \sup_{\phi} e_t(\phi) \\ 1(\bar{A}_{3,t-1}(\phi^*, \nu)) + 1(\bar{A}_{4,t-1}(\phi^*, \nu)) & \geq \sup_{\phi \in \Theta | \rho(\phi, \phi^*) < \nu} |1(A_{3,t-1}(\phi, \phi^*)) + 1(A_{4,t-1}(\phi, \phi^*))| \\ \bar{v}_t(\phi^*, \nu) & \geq \sup_{\phi \in \Theta | \rho(\phi, \phi^*) < \nu} v_t(\phi, \phi^*) \\ \bar{v}_t(\beta^*, \nu) & \geq \sup_{\beta \in \Theta_\beta | \rho(\beta, \beta^*) < \nu} v_t(\beta, \beta^*). \end{aligned}$$

Como en los casos anteriores, en la prueba usaremos los siguientes lemas que establecen las propiedades de dependencia y momentos de los procesos que usaremos como cotas.

Lema 3.6 *Bajo los supuestos G.0 y ASG.1 se cumple que*

$$(a) \ E(\bar{e}_t^{2\gamma}) \leq K_e < \infty$$

- (b) $E \left(\bar{v}_t^{2\gamma} (\beta^*, \nu) \right) \leq \nu^{2\gamma} K_v < \infty$, si $r^0 = 0$
- (c) $E \left(\bar{v}_t^{2\gamma} (\phi^*, \nu) \right) \leq \nu K_v < \infty$, si $|r^0| > 0$
- (d) $E \left[1 \left(\bar{A}_{3,t-1} (\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1} (\phi^*, \nu) \right) \right] \leq \nu K_A < \infty$.

Prueba del Lema 3.6. La prueba de (a), (c) y (d) es similar a las del lema 3.3. El apartado (b), se obtiene directamente de la definición de $\bar{v}_t(\beta^*, \nu)$ cuando $r^0 = 0$ y conocido. ■

Lema 3.7 *Bajo los supuestos G.0 y ASG.1,*

- (a) $e_t(\phi)$ es $L_{2\gamma}$ -acotado y L_2 -NED de tamaño $-\infty$ en ε_t , con constante $\varrho^{1/2}(\phi, \phi^0) K_e$
- (b) \bar{e}_t es $L_{2\gamma}$ -acotado y L_2 -NED de tamaño $-\infty$ en ε_t , con constante $d_{\bar{e},t}$ finita.

Prueba del Lema 3.7. La primera parte de ambos apartados se obtiene del lema 3.6. Como en el caso del tamaño, se puede probar que x_t es L_2 -NED de tamaño $-\infty$ en ε_t con $v_m = \rho^m$ y $d_t = \|x_t\|_2$. Definimos,

$$e_t^m = \varepsilon_t - \theta_{t-1}^0 \varepsilon_{t-1} + (\rho^0 - \rho) x_{t-1}^{m-1} + (\mu^0 - \mu) + \theta_{t-1} (e_{t-1}^{m-1}) e_{t-1}^{m-1},$$

donde $\theta_{t-1} (e_{t-1}^{m-1}) = \theta_1 1(e_{t-1}^{m-1} > r) + \theta_2 1(e_{t-1}^{m-1} < r)$ y $e_{t-m+1}^1 = \varepsilon_{t-m+1}$ tal que e_t^m es \mathfrak{S}_{t-m}^{t+m} -medible. Así,

$$\begin{aligned} \|e_t - e_t^m\|_1 &\leq \|(\rho^0 - \rho) (x_{t-1} - x_{t-1}^{m-1})\|_1 + \max_i |\theta_i| \|e_{t-1} - e_{t-1}^{m-1}\|_1 + \\ &\quad \partial r \|E[1(|e_{t-1}| > r, |e_{t-1}^{m-1}| < r) + 1(|e_{t-1}| < r, |e_{t-1}^{m-1}| > r)]\|_1, \end{aligned}$$

y dado que $(e_{t-1} - e_{t-1}^{m-1})$ es \mathfrak{S}_{t-2} -medible, obtenemos que

$$\|e_t - e_t^m\|_1 \leq (\rho^0 - \rho) \rho^{m-1} \|x_t\|_1 + \max_i |\theta_i| \|e_{t-1} - e_{t-1}^{m-1}\|_1 + \partial r f^* \|e_{t-1} - e_{t-1}^{m-1}\|_1.$$

Usando el supuesto ASZ.1,

$$\begin{aligned}
\|e_t - e_t^m\|_1 &\leq (\rho^0 - \rho) \rho^{m-1} \|x_t\|_2 + \lambda_1(\phi) \|e_{t-1} - e_{t-1}^{m-1}\|_1 \leq \\
&\leq (\rho^0 - \rho) \rho^{m-1} \|x_t\|_1 \left[1 + \sum_{j=1}^{m-2} \lambda_1^j \rho^{-j} \right] + \lambda_1^j(\phi) \|e_{t-m+1} - e_{t-m+1}^1\|_1 \\
&\leq m \max(\rho, \lambda_1(\phi))^{m-1} [(\rho^0 - \rho) \|x_t\|_1 + \|v_{t-m+1}(\phi, \phi^0)\|_1] \\
&\leq m \lambda_\rho^{m-1} d_{e,t,1}.
\end{aligned}$$

donde $\lambda_\rho = \max(\rho, \lambda_1) < 1$ y $d_{e,t,1} = [(\rho^0 - \rho) \|x_t\|_1 + \|v_{t-m+1}(\phi, \phi^0)\|_1]$. Como en los casos anteriores, para la norma de orden dos, obtenemos que

$$\|e_t - e_t^m\|_2 \leq m^{3/2} \lambda_\rho^{\frac{m-1}{2}} \left[(\rho^0 - \rho) \|x_t\|_2 + \partial r(f^* d_{e,t,1})^{1/2} + \|v_{t-m+1}(\phi, \phi^0)\|_2 \right].$$

Definiendo $v_t = m^{3/2} \lambda_\rho^{\frac{m-1}{2}}$ y $d_{e,t} = [(\rho^0 - \rho) \|x_t\|_2 + \partial r(f^* d_{e,t,1})^{1/2} + \|v_{t-m+1}(\phi, \phi^0)\|_2]$, queda probado el apartado (a). Para \bar{e}_t la prueba es similar. ■

Lema 3.8 Para el caso $|r^0| > 0$, bajo los supuestos G.0 y A.SZ.1, $\bar{v}_t(\phi^*, \nu)$ es una L_2 -mixingala en ε_t de tamaño $-1/2$ y constante $\max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v^{1/2}$ con $\bar{K}_v < \infty$. Por lo tanto,

$$E \left[\max_{1 \leq j \leq T} \left(\sum_{t=1}^j \chi_t \right)^2 \right] \leq K \sum_{t=1}^T \max(\varrho(\phi^*, \phi^0), \nu) \bar{K}_v,$$

para $\chi_t = |\varepsilon_t| \bar{v}_t(\phi^*, \nu) - E(|\varepsilon_t| \bar{v}_t(\phi^*, \nu))$.

Prueba del Lema 3.8. Para probar que $\bar{v}_t(\phi^*, \nu)$ es L_2 -mixingala, como en el caso del tamaño, primero probamos que es una secuencia L_2 -NED en ε_t con constante $\max(\rho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v$. Definimos la siguiente variable \mathfrak{S}_{t-m}^{t+m} -medible,

$$\begin{aligned}
\bar{v}_t^m(\phi^*, \nu) &= \nu \left[|x_{t-1}^{m-1}| + 3 + (r^* + \nu) + 3\bar{e}_{t-1}^{m-1} \right] + \max_i |\theta_i^*| \bar{v}_{t-1}^{m-1}(\phi^*, \nu) \\
&\quad + \partial^* r^* \left[1 \left(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu) \right) \right],
\end{aligned}$$

donde,

$$\begin{aligned} & 1 \left(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu) \right) = \\ & = 1 \left[r^* - \nu - v_{t-1}^{m-1}(\phi^*, \phi^0) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu) < \varepsilon_{t-1} < r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) \right] + \\ & + 1 \left[-r^* - v_{t-1}^{m-1}(\phi^*, \phi^0) < \varepsilon_{t-1} < -r^* + \nu - v_{t-1}^{m-1}(\phi^*, \phi^0) + \bar{v}_{t-1}^{m-1}(\phi^*, \nu) \right]. \end{aligned}$$

De forma similar al lema 3.5, se obtiene que,

$$\begin{aligned} \|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_1 & \leq \nu \left[\|x_{t-1} - x_{t-1}^{m-1}\|_1 + 3 \|\bar{e}_{t-1} - \bar{e}_{t-1}^{m-1}\|_1 \right] + \\ & + \max_i |\theta_i^*| \|\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_1 + \\ & + \delta^* r^* \left\| 1 \left(\bar{A}_{3,t-1}(\phi^*, \nu) \right) - 1 \left(\bar{A}_{3,t-1}^{m-1}(\phi^*, \nu) \right) + 1 \left(\bar{A}_{4,t-1}(\phi^*, \nu) \right) - 1 \left(\bar{A}_{4,t-1}^{m-1}(\phi^*, \nu) \right) \right\|_1, \end{aligned}$$

y que $v_t(\phi^*, \phi^0)$ es $L_2 - NED$, en ε_t , de tamaño $-\infty$ y constante $d_{v,t}^0 = \rho^{1/2}(\phi^*, \phi^0) K_v$.

Además,

$$\begin{aligned} \left\| \sum_{i=3}^4 1 \left(\bar{A}_{i,t-1}(\phi^*, \nu) \right) - 1 \left(\bar{A}_{i,t-1}^{m-1}(\phi^*, \nu) \right) \right\|_1 & \leq f^* \|\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_1 + \\ & + f^* \|v_{t-1}(\phi^*, \phi^0) - v_{t-1}^{m-1}(\phi^*, \phi^0)\|_1. \end{aligned}$$

Por lo tanto, agrupando resultados,

$$\begin{aligned} \|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_1 & \leq \nu \left[\rho^{m-1} d_{x,t} + 3m\lambda_\rho^{m-1} d_{\bar{e},t} \right] + \\ & + \left(\max_i |\theta_i^*| + f^* \delta^* r^* \right) \|\bar{v}_{t-1}(\phi^*, \nu) - \bar{v}_{t-1}^{m-1}(\phi^*, \nu)\|_1 + m\lambda_\rho^{m-1} d_{e,t}^1 f^* \delta^* r^* \\ & \leq m^2 \lambda_\rho^{\frac{m-1}{2}} \max(\rho(\phi^*, \phi^0), \nu) K_v. \end{aligned}$$

Siguiendo los mismos pasos que en el lema 3.5, se concluye que,

$$\|\bar{v}_t(\phi^*, \nu) - \bar{v}_t^m(\phi^*, \nu)\|_2 \leq m^2 \lambda_\rho^{\frac{m-1}{4}} \max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v^{1/2},$$

que implica que $\bar{v}_t(\phi^*, \nu)$ es una secuencia $L_2 - NED$ en ε_t de tamaño $-\infty$ y constante $d_{v,t} = \max(\varrho^{1/2}(\phi^*, \phi^0), \nu^{1/2}) \bar{K}_v^{1/2}$, con lo que se deriva el resultado del lema por el

Teorema 2.10 y 2.12 del Anexo 1 en el Capítulo 2 y dado que ε_t es *iid*. ■

Lema 3.9 Para el caso $r^0 = 0$, bajo los supuestos G.0 y A.SZ.1, $\bar{v}_t(\beta^*, \nu)$ es una L_2 -mixingala en ε_t de tamaño $-1/2$ y constante $\nu \bar{K}_v^{1/2}$ con $\bar{K}_v < \infty$. Por lo tanto,

$$E \left[\max_{1 \leq j \leq T} \left(\sum_{t=1}^j \chi_t \right)^2 \right] \leq K \sum_{t=1}^T \nu^2 \bar{K}_v,$$

para $\chi_t = |\varepsilon_t| \bar{v}_t(\beta^*, \nu) - E(|\varepsilon_t| \bar{v}_t(\beta^*, \nu))$.

Prueba del Lema 3.9. Usando la definición de $\bar{v}_t(\beta^*, \nu)$ para el caso $r^0 = 0$, definimos la siguiente variable \mathfrak{S}_{t-m}^{t+m} -measurable,

$$\bar{v}_t^m(\beta^*, \nu) = \nu [|x_{t-1}|^{m-1} + 1 + 3\bar{e}_{t-1}^{m-1}] + \max_i |\theta_i^*| \bar{v}_{t-1}^{m-1}(\beta^*, \nu).$$

En este caso, calculamos directamente la norma 2, esto es

$$\begin{aligned} \|\bar{v}_t(\beta^*, \nu) - \bar{v}_t^m(\beta^*, \nu)\|_2 &\leq \nu [\| |x_{t-1}| - |x_{t-1}|^{m-1} \|_2 + 3 \|\bar{e}_{t-1} - \bar{e}_{t-1}^{m-1}\|_2] + \\ &\quad + \max_i |\theta_i^*| \|\bar{v}_{t-1}(\beta^*, \nu) - \bar{v}_{t-1}^{m-1}(\beta^*, \nu)\|_2 \\ &\leq m^2 \lambda_\rho^{\frac{m-1}{2}} \nu \bar{K}_v^{1/2}. \end{aligned}$$

De esta forma, $\bar{v}_t(\beta^*, \nu)$ será L_2 -NED en ε_t de tamaño $-\infty$ y constante, en este caso, $d_{v,t} = \nu \bar{K}_v^{1/2}$, lo cual implica que será L_2 -mixingala de tamaño $-1/2$ y constante $\nu \bar{K}_v^{1/2}$.

■

Dados estos resultados, las pruebas de las proposiciones 2.15 y 2.16, de las que derivamos la consistencia de los estimadores se obtienen de forma directa para ambos casos, $r^0 = 0$ conocido y $|r^0| > 0$, por lo que, por motivos de espacio no las reproducimos aquí. A continuación, probamos el Teorema 3.2 que establece la tasa de convergencia.

Prueba del Teorema 3.2

Empezamos por el caso $r^0 = 0$ conocido, en el que los estimadores presentan una tasa de convergencia $T^{-1/2}$.

Caso $r^0 = 0$.

Definimos $B_{\Delta, \epsilon} = \left\{ \beta \in \Theta \mid \sqrt{\frac{\Delta}{T}} \leq |\beta_i - \beta_i^0| \leq \epsilon < 1, \text{ para } i = 1, 2, 3, 4 \right\}$. El resultado del Teorema, se deriva de las proposiciones 3.9 y 3.10, que presentamos a continuación.

Proposición 3.9 *Para todo $\eta, \varepsilon > 0 \exists \Delta, T_\varepsilon < \infty$ tal que para todo $T > T_\varepsilon$,*

$$P \left[\sup_{\beta \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t(\beta, \beta^0)}{T \varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon$$

Proposición 3.10 *Para todo $\varepsilon > 0 \exists \Delta, T_\varepsilon < \infty$ y $\eta > 0$ tal que para todo $T > T_\varepsilon$,*

$$P \left[\sup_{\beta \in B_{\Delta, \epsilon}} \frac{\sum_{t=1}^T v_t^2(\beta, \beta^0)}{T \varrho^2(\beta, \beta^0)} > \eta \right] \geq 1 - \varepsilon$$

Prueba de la Proposición 3.9. Definimos,

$$e_t^*(\beta) = x_t - \rho x_{t-1} - \mu + \theta_1 1(\varepsilon_{t-1} > 0) e_{t-1}^*(\beta) + \theta_2 1(\varepsilon_{t-1} < 0) e_{t-1}^*(\beta)$$

$$v_t^*(\beta) = e_t(\beta) - e_t^*(\beta) = \theta(\varepsilon_{t-1}) v_{t-1}^*(\beta) + (\theta_1 - \theta_2) e_{t-1} 1(A_{3,t-1}(\beta, \beta^0)) \\ + (\theta_2 - \theta_1) e_{t-1} 1(A_{4,t-1}(\beta, \beta^0))$$

$$\bar{s}_t(\beta, \nu) = \nu v_{t-1}^*(\beta) + (1 - \delta) \bar{s}_{t-1}(\beta, \nu) + 2\nu v_{t-1}(\beta, \beta^0) + \\ + 2\bar{v}_{t-1}(\beta, \nu) \sum_{i=3}^4 1(A_{i,t-1}(\beta, \beta^0)) + 2v_{t-1}(\beta, \beta^0) \sum_{i=3}^4 1(\bar{A}_{i,t-1}(\beta, \nu)),$$

donde $\bar{s}_t(\beta, \nu) \geq \sup_{\beta \in B_{\Delta, \epsilon} / \varrho(\beta, \beta^*) \leq \nu} v_t^*(\beta) - v_t^*(\beta^*)$. Usando la LIE y el lema 3.6, es fácil probar que para $|\varrho(\beta, \beta^0)| < 1$,

$$\|v_t^*(\beta)\|_2 \leq \varrho(\beta, \beta^0) K_v$$

$$\|\bar{s}_t(\beta, \nu)\|_2 \leq \nu \varrho(\beta, \beta^0) K_s.$$

Además, usando el lema 3.9, y siguiendo los mismos pasos que en dicho lema, se obtiene que $v_t^*(\beta)$ y $\bar{s}_t(\beta, \nu)$ son L_2 - *mixingales* de tamaño $-1/2$ y constantes $\|v_t^*(\beta)\|_2$ y $\|\bar{s}_t(\beta, \nu)\|_2$ respectivamente. Para probar la proposición 3.9, será suficiente probar que

para todo $\eta, \varepsilon > 0 \exists \Delta, T_\varepsilon < \infty$ tal que para todo $T > T_\varepsilon$,

$$P \left[\sup_{\beta \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t (e_t^*(\beta) - \varepsilon_t)}{T \varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon \quad (3.23)$$

$$P \left[\sup_{\beta \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t^*(\beta)}{T \varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon. \quad (3.24)$$

La desigualdad (3.23) se obtiene fácilmente teniendo en cuenta que $(e_t^*(\beta) - \varepsilon_t)$ es dos veces diferenciable respecto a β , (ver la prueba del Teorema 2.2 en el Anexo 2 del Capítulo anterior). Para la desigualdad (3.24), definimos la siguiente partición, $B_{\varepsilon, j_l}^l(\Delta, \beta_l) = \left\{ \beta_l \in B_{\Delta, \varepsilon} \left/ \frac{\beta_l - \beta_l^{j_l}}{\beta_l^{j_l} - \beta_l^0} \leq (b-1) \text{ para } l = 1, 2, 3, 4 \right. \right\}$ con $\beta_l^{j_l} = \beta_l^0 \pm \sqrt{\frac{\Delta}{T}} b^{j_l}$ para $l = 1, \dots, 4$ y $b > 1$. Por lo tanto, definiendo, $\ell_j = (j_1, j_2, j_3, j_4)$, para $j = 1, \dots, J$, es directo probar que $B_{\Delta, \varepsilon} \subset \cup_{\ell_j=1}^J B_{\varepsilon, \ell_j}^\Delta(\beta)$, con $B_{\varepsilon, \ell_j}^\Delta(\beta) = \prod_{l=1}^4 B_{\varepsilon, j_l}^l(\Delta, \beta_l)$. En este caso, con $\beta^{\ell_j} = (\beta_1^{j_1}, \beta_2^{j_2}, \beta_3^{j_3}, \beta_4^{j_4})$, $\sup_{\phi \in B_{\varepsilon, \ell_j}^\Delta(\beta)} \varrho(\beta^{\ell_j}, \beta) = \nu_\Delta(\ell_j) = \varrho(\beta^{\ell_j}, \beta^0)(b-1)$. Siguiendo los pasos de casos anteriores, conseguimos las siguientes tres desigualdades,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T \varepsilon_t v_t^*(\beta^{\ell_j})}{T \varrho^2(\beta^{\ell_j}, \beta^0)} \right| > \eta \right] &\leq \frac{4\sigma_\varepsilon^2 K_v}{\Delta \eta^2} \sum_{j=1}^\infty \frac{j^3}{b^j}, \\ P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T |\varepsilon_t| \bar{s}_t(\beta^{\ell_j}, \nu_\Delta(\ell_j)) - E[|\varepsilon_t| \bar{s}_t(\beta^{\ell_j}, \nu_\Delta(\ell_j))]}{T \varrho^2(\beta^{\ell_j}, \beta^0)} \right| > \eta \right] &\leq \frac{4\sigma_\varepsilon^2 K_v}{\Delta \eta^2} \sum_{j=1}^\infty \frac{j^3}{b^j}, \\ \max_{\ell_j} \frac{E[|\varepsilon_t| \bar{s}_t(\beta^{\ell_j}, \nu_\Delta(\ell_j))]}{\varrho^2(\beta^{\ell_j}, \beta^0)} &\leq \sigma_\varepsilon K_v^{1/2} (b-1). \end{aligned}$$

Con estas desigualdades, la prueba de la proposición se obtiene de forma directa.

■

Prueba de la Proposición 3.10. Dado que r^0 es conocido, definimos

$$W_{k(\iota)}(\beta, \beta^0) = v_{k(\iota)-1}^2(\beta, \beta^0) + v_{k(\iota)}^2(\beta, \beta^0) \quad \text{con } k(\iota) = 2\iota.$$

Sea $\delta_{\phi, l}^0 = \phi_l - \phi_l^0$, para $l = 1, 2, 3, 4$, entonces $v_{k(\iota)}(\beta, \beta^0)$ puede escribirse como,

$$v_{k(\iota)}(\beta, \beta^0) = \delta_{\phi, 3}^0 x_{k(\iota)-1} + \delta_{\phi, 4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1} + \theta_{k(\iota)-1} v_{k(\iota)-1}(\beta, \beta^0),$$

con $\theta_{k(\iota)-1} = \theta_{k(\iota)-1}(\varepsilon_{k(\iota)-1}(\beta))$ y $\theta_{k(\iota)-1}^0 = \theta_{k(\iota)-1}^0(\varepsilon_{k(\iota)-1})$.

En el caso del tamaño o $|r^0| > 0$, el término que garantiza la T consistencia es $[(\delta^0 - \epsilon)(r^0 - \epsilon)]^2 [1(A_{3,t-1}(\phi, \phi^0)) + 1(A_{4,t-1}(\phi, \phi^0))]$, que, sin embargo, para $r^0 = 0$ será tan pequeño como queramos. Por lo tanto, nos centramos en el error cometido en los sucesos $A_{1,t-1}(\beta, \beta^0)$ y $A_{2,t-1}(\beta, \beta^0)$. Así, condicionando a los sucesos

$$|v_{k(\iota)-1}^2(\beta, \beta^0)| \leq \frac{1}{2} \min \{ |\delta_{\phi,3}^0 x_{k(\iota)-1} + \delta_{\phi,4}^0 + (\theta_{k(\iota)-1} - \theta_{k(\iota)-1}^0) \varepsilon_{k(\iota)-1}|, \epsilon^2 M \},$$

con $\bar{r} = \epsilon^2 M$, obtenemos que

$$\begin{aligned} W_{k(\iota)}(\beta, \beta^0) &\geq \min(\epsilon M, \vartheta_{k(\iota),1}^2(\delta_\phi^0)) \\ \vartheta_{k(\iota),1}^2(\delta_\phi^0) &= \frac{1}{2} 1(\varepsilon_{k(\iota)-1} > \bar{r}) |\delta_{\phi,3}^0 x_{k(\iota)-1} + \delta_{\phi,4}^0 + \delta_{\phi,1}^0 \varepsilon_{k(\iota)-1}|^2 + \\ &\quad \frac{1}{2} 1(\varepsilon_{k(\iota)-1} < \underline{r}) |\delta_{\phi,3}^0 x_{k(\iota)-1} + \delta_{\phi,4}^0 + \delta_{\phi,2}^0 \varepsilon_{k(\iota)-1}|^2 \\ &= \max_l (\delta_{\phi,l}^0)^2 [c' M_{k(\iota)-1}(\underline{r}, \bar{r}) c], \end{aligned}$$

con $c \in [-1, 1]^6$ y $\max_i c_i = 1$. Por el supuesto ASZ.2,

$$\inf_c E [c' M_{k(\iota)-1}(\underline{r}, \bar{r}) c] \geq a > 0.$$

Definiendo $\bar{M}_{k(\iota)-1} = 2(|x_{k(\iota)-1}| + 1 + |\varepsilon_{k(\iota)-1}|)^2$, y dado que $E(\bar{M}_{k(\iota)-1}^2) < \infty$, $\exists M < \infty$ tal que $\inf_c E [c' M_{k(\iota)-1}(\underline{r}, \bar{r}) c 1(|\bar{M}_{k(\iota)-1}| < M)] > a/2 > 0$. Ahora, por la consistencia de los estimadores, existirá T suficientemente grande de forma que, $\max_l \hat{\delta}_{\phi,l}^0 < \epsilon = (\bar{r}/M)^{1/2}$, con probabilidad tan grande como queramos, con lo que

$$W_{k(\iota)}(\beta, \beta^0) \geq \max_l (\delta_{\phi,l}^0)^2 [c' M_{k(\iota)-1}(\bar{r}, \underline{r}) c] 1(|\bar{M}_{k(\iota)-1}| < M),$$

con la misma probabilidad. Como en el caso de la proposición 3.8, obtenemos que $M_{k(\iota)-1}(\bar{r}, \underline{r}) 1(|\bar{M}_{k(\iota)-1}| < M)$ es L_2 - *mixigala* de tamaño $-1/2$ y constante finita. Por lo tanto, para $\forall \eta > 0$

$$\lim_{T \rightarrow \infty} P \left[\sup_c \left| \frac{\sum_{\iota=1}^{\lfloor \frac{T}{2} \rfloor} c' [\widetilde{M}_{k(\iota)-1}(\bar{r}, \underline{r}) - E(\widetilde{M}_{k(\iota)-1}(\bar{r}, \underline{r}))] c}{T} \right| > \eta \right] = 0,$$

con $\widetilde{M}_{k(l)-1}(\bar{r}, \underline{r}) = M_{k(l)-1}(\bar{r}, \underline{r}) 1(|\bar{M}_{k(l)-1}| < M)$. De esta manera obtenemos que

$$\lim_{T \rightarrow \infty} P \left[\inf_{\{\beta \in B_{\Delta, \epsilon}\}} \frac{\sum_{t=1}^T (v_t(\beta, \beta^0))^2}{T \varrho^2(\phi, \phi^0)} > a/2 - \eta \right] = 1$$

con $a/2 - \eta > 0$ para $\eta > 0$ suficientemente pequeño. ■

Caso $|r^0| > 0$.

La prueba del Teorema 3.2 para este caso, se obtiene de la misma forma que la prueba del Teorema 3.1, por lo que no la desarrollaremos aquí.

3.9.2. Inferencia.

Inferencia con $r^0 = 0$ conocido.

Para este caso definimos,

$$\begin{aligned} d_t(\beta) &= (\rho^0 - \rho) x_{t-1} + (\mu^0 - \mu) + (\theta_1 - \theta_1^0) \varepsilon_{t-1} 1(\varepsilon_{t-1} > 0) + (\theta_2 - \theta_2^0) \varepsilon_{t-1} 1(\varepsilon_{t-1} < 0) + \\ &\quad + \theta_{t-1}^0 d_{t-1}(\beta) \\ &= (\rho^0 - \rho) \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 x_{t-1-j} + (\mu^0 - \mu) \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 + (\theta_1 - \theta_1^0) \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^+ + \\ &\quad + (\theta_2 - \theta_2^0) \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}^0 \varepsilon_{t-1-j}^- \\ &= \sum_{i=1}^4 (\beta_i - \beta_i^0) h_t^{(i)}, \end{aligned}$$

con $\varepsilon_{t-j}^+ = \varepsilon_{t-1} 1(\varepsilon_{t-1} > 0)$ y $\varepsilon_{t-j}^- = \varepsilon_{t-1} - \varepsilon_{t-j}^+$.

Primero probaremos que

$$Q_T(\widehat{\beta}) - Q_T(\beta^0) = \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\widehat{\beta}_i - \beta_i^0) H_t^{i,l} (\widehat{\beta}_l - \beta_l^0) + \sum_{i=1}^4 \sum_{t=1}^T 2 (\widehat{\beta}_i - \beta_i^0) \varepsilon_t h_t^{(i)} + o_p(1),$$

con el $o_p(1)$ uniforme sobre β y $H_t^{i,l} = h_t^{(i)} h_t^{(l)}$. Esta igualdad, se deriva de las siguientes dos proposiciones,

Proposición 3.11 *Bajo los supuestos G.0, G.2, G.4, ASG.1 y ASG.2,*

$$\sum_{t=1}^T \varepsilon_t \left(v_t \left(\widehat{\beta}, \beta^0 \right) - d_t \left(\widehat{\beta} \right) \right) = o_p(1).$$

Proposición 3.12 *Bajo los supuestos G.0, G.2 y ASG.1,*

$$\sum_{t=1}^T \left(v_t^2 \left(\widehat{\beta}, \beta^0 \right) - d_t^2 \left(\widehat{\beta} \right) \right) = o_p(1).$$

Prueba de la Proposición 3.11. El resultado se sigue si para todo $\eta > 0$ y $\varepsilon > 0$, $\exists T_\varepsilon$ tal que para todo $T > T_\varepsilon$

$$P \left[\left| \sum_{t=1}^T \varepsilon_t \left(v_t \left(\widehat{\beta}, \beta^0 \right) - d_t \left(\widehat{\beta} \right) \right) \right| > \eta \right] < \varepsilon.$$

Del Teorema 3.2, sabemos que $T^{-1/2} \left(\widehat{\beta}_i - \beta_i^0 \right) = O_p(1)$ para $i = 1, 2, 3, 4$. Por lo tanto, para $\varepsilon > 0$ dado, podemos escoger Δ suficientemente grande tal que $P \left[\max_i \left| \widehat{\beta}_i - \beta_i^0 \right| > T^{-1/2} \Delta \right] < \varepsilon/2$. Definiendo $D_\Delta = [-T^{-1/2} \Delta, T^{-1/2} \Delta]^4$, el resultado de la proposición se obtendrá, si para todo $\eta > 0$ y $\varepsilon > 0$, $\exists T_\varepsilon$ tal que para todo $T > T_\varepsilon$

$$P \left[\sup_{\beta \in D_\Delta} \left| \sum_{t=1}^T \varepsilon_t \left(v_t \left(\beta, \beta^0 \right) - d_t \left(\beta \right) \right) \right| > \eta \right] < \varepsilon/2.$$

De la definición de $v_t(\beta, \beta^0)$ y $d_t(\beta)$, se obtiene que

$$\begin{aligned} v_t(\beta, \beta^0) - d_t(\beta) &= \theta_{t-1}^0 \left(v_{t-1}(\beta, \beta^0) - d_{t-1}(\beta) \right) + \\ &+ \left[(\theta_1 - \theta_2) 1(A_{3,t-1}(\beta, \beta^0)) + (\theta_2 - \theta_1) 1(A_{4,t-1}(\beta, \beta^0)) \right] \varepsilon_{t-1} + \\ &+ \left[(\theta_1 - \theta_1^0) 1(\varepsilon_{t-1} > 0) + (\theta_2 - \theta_2^0) 1(\varepsilon_{t-1} < 0) \right] v_{t-1}(\beta, \beta^0). \end{aligned}$$

Ahora, definiendo

$$\begin{aligned} \bar{c}_t(\beta, \nu) &= \theta_{t-1}^0 \bar{c}_{t-1}(\beta, \nu) + 2\nu \left[1(A_{3,t-1}(\beta, \beta^0)) + 1(A_{4,t-1}(\beta, \beta^0)) \right] \varepsilon_{t-1} + \\ &+ 4 \left[1(\bar{A}_{4,t-1}(\beta, \nu)) + 1(\bar{A}_{3,t-1}(\beta, \nu)) \right] \varepsilon_{t-1} + 2 \left[\varrho(\beta, \beta^0) + v \right] \bar{v}_{t-1}(\beta, \nu) + \\ &+ 2\nu v_{t-1}(\beta, \beta^0), \end{aligned}$$

se puede probar que $\sup_{\phi/\rho(\phi^*, \phi) \leq \nu} |(v_t(\beta, \beta^0) - d_t(\beta)) - (v_t(\beta^*, \beta^0) - d_t(\beta^*))| \leq \bar{c}_t(\beta^*, \nu)$. A través de la LIE y los supuestos de la proposición, y dado que $\bar{v}_{t-1}(\beta, \nu)$ y $v_{t-1}(\beta, \beta^0)$ son \mathfrak{S}_{t-2} -medibles, se tiene que

$$E(\varepsilon_t^2 \bar{c}_t^2(\beta, \nu)) \leq \max(\rho^2(\beta, \beta^0) \nu, \nu^3) K_s$$

$$E(|\varepsilon_t| \bar{c}_t(\beta, \nu)) \leq \max(\rho(\beta, \beta^0) \nu, \nu^2) K_s.$$

Además, usando el Teorema 2.12 y de la forma habitual, se tiene que $|\varepsilon_t| \bar{c}_t(\beta, \nu)$ es una secuencia L_2 -mixingala en ε_t de tamaño $-1/2$ con constante $\max(\rho(\beta, \beta^0) \nu^{1/2}, \nu^{3/2}) K_s$, con $K_s < \infty$.

Como en casos anteriores, definimos la siguiente partición para D_Δ , $B_{\ell_j}^\Delta(\beta) = \prod_{l=1}^4 B_{j_l}^l(\Delta, \beta_l)$, donde $B_{j_l}^l(\Delta, \beta_l) = \{\beta_l \in D_\Delta / |\beta_l^{j_l} - \beta_l| \leq T^{-1/2} \Delta \nu\}$ con $\beta_l^{j_l} = \beta_l^0 \pm T^{-1/2} \Delta \nu j_l$ para $l = 1, \dots, 4$. Entonces, definiendo, $\ell_j = (j_1, j_2, j_3, j_4)$, para $j = 1, \dots, J$, con $J = 1/\nu$, es directo probar que $D_\Delta \subset \cup_{\ell_j=1}^J B_{\ell_j}^\Delta(\beta)$. Para este caso y con $\beta^{\ell_j} = (\beta_1^{j_1}, \beta_2^{j_2}, \beta_3^{j_3}, \beta_4^{j_4})$, $\sup_{\beta \in B_{\ell_j}^\Delta(\beta)} \varrho(\beta^{\ell_j}, \beta) = T^{-1/2} \Delta \nu$. Ahora,

$$\begin{aligned} \sup_{\beta \in D_\Delta} \left| \sum_{t=1}^T \varepsilon_t (v_t(\beta, \beta^0) - d_t(\beta)) \right| &\leq \max_{\ell_j} \left| \sum_{t=1}^T \varepsilon_t (v_t(\beta^{\ell_j}, \beta^0) - d_t(\beta^{\ell_j})) \right| + \\ &+ \max_{\ell_j} \left| \sum_{t=1}^T |\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) - E \left(|\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) \right) \right| + \\ &+ \max_{\ell_j} \left| \sum_{t=1}^T E \left(|\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) \right) \right|. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \sum_{t=1}^T \varepsilon_t (v_t(\beta^{\ell_j}, \beta^0) - d_t(\beta^{\ell_j})) \right| > \eta/2 \right] &\leq \sum_{j=1}^J \frac{j^3 \nu^2 K_s}{\eta^2 T^{1/2}} \leq \frac{K_s}{\eta^2 \nu^2 T^{1/2}} \xrightarrow{T \rightarrow \infty} 0, \\ P \left[\max_{\ell_j} \left| \sum_{t=1}^T |\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) - E \left(|\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) \right) \right| > \eta/4 \right] &\leq \frac{K_s}{\eta^2 \nu^4 T^{1/2}} \xrightarrow{T \rightarrow \infty} 0, \\ \max_{\ell_j} \left| \sum_{t=1}^T E \left(|\varepsilon_t| \bar{c}_t \left(\beta^{\ell_j}, \frac{\Delta}{T^{1/2}} \nu \right) \right) \right| &\leq \nu K_s < \eta/4, \end{aligned}$$

para ν suficientemente pequeño. La primera desigualdad se sigue de

$E \left[\left(v_t(\beta^{\ell_j}, \phi^0) - d_t(\beta^{\ell_j}) \right)^2 \right] \leq E \left[\bar{c}_t(\beta^0, \frac{\Delta}{T^{1/2}} \nu \max \ell_j) \right]$. Con estos resultados, la proposición 3.11 queda probada. ■

Prueba de la Proposición 3.12. Para este caso,

$$\begin{aligned} \sum_{t=1}^T \left(v_t^2(\hat{\beta}, \beta^0) - d_t^2(\hat{\beta}) \right) &= \sum_{t=1}^T \left(v_t(\hat{\beta}, \beta^0) - d_t(\hat{\beta}) \right) v_t(\hat{\beta}, \beta^0) + \\ &+ \sum_{t=1}^T \left(v_t(\hat{\beta}, \beta^0) - d_t(\hat{\beta}) \right) d_t(\hat{\beta}). \end{aligned}$$

Por lo tanto, la proposición se sigue de

$$\begin{aligned} P \left[\sup_{\beta \in D_\Delta} \left| \sum_{t=1}^T (v_t(\beta, \beta^0) - d_t(\beta)) v_t(\beta, \beta^0) \right| > \eta/2 \right] &< \varepsilon/2 \\ P \left[\sup_{\beta \in D_\Delta} \left| \sum_{t=1}^T (v_t(\beta, \beta^0) - d_t(\beta)) d_t(\beta) \right| > \eta/2 \right] &< \varepsilon/2. \end{aligned}$$

Dado que $\varepsilon^0 = 0$, $\bar{c}_t(\beta^0, \frac{\Delta}{T^{1/2}}) \geq v_t(\beta, \beta^0) - d_t(\beta)$, $\|\bar{c}_t^2(\beta^0, \frac{\Delta}{T^{1/2}})\|_2 \leq (\frac{\Delta}{T})^{3/4} K_c$ y $\|\bar{v}_t(\beta^0, \frac{\Delta}{T^{1/2}})\|_2 \leq (\frac{\Delta}{T})^{1/2} \bar{K}_v$ se tiene que,

$$\begin{aligned} P \left[\sup_{\beta \in D_\Delta} \left| \sum_{t=1}^T (v_t(\beta, \beta^0) - d_t(\beta)) v_t(\beta, \beta^0) \right| > \eta/2 \right] &\leq \\ &\leq P \left[\left| \sum_{t=1}^T \left(\bar{c}_t\left(\beta^0, \frac{\Delta}{T^{1/2}}\right) \right) \bar{v}_t\left(\beta^0, \frac{\Delta}{T^{1/2}}\right) \right| > \eta/2 \right] \\ &\leq \frac{T \|\bar{c}_t(\beta^0, \frac{\Delta}{T^{1/2}})\|_2 \|\bar{v}_t(\beta^0, \frac{\Delta}{T^{1/2}})\|_2}{\eta} \leq \frac{K_{sv}}{\eta T^{1/4}} \xrightarrow{T \rightarrow \infty} 0. \end{aligned}$$

El mismo resultado puede obtenerse para la otra probabilidad, teniendo en cuenta que $|d_t(\beta)| \leq \frac{\Delta}{T^{1/2}} \sum_{l=1}^4 \left| h_t^{(i)}(\beta^0) \right|$ con $\left\| h_t^{(i)}(\beta^0) \right\|_2 \leq \bar{K}_h < \infty$. Con esto, queda probada la proposición 3.12. ■

Prueba del Teorema 3.3. Por las proposiciones 3.11 y 3.12 obtenemos que

$$\begin{aligned} Q_T(\hat{\beta}) - Q_T(\beta^0) &= \sum_{t=1}^T d_t^2(\hat{\beta}) + \sum_{t=1}^T 2\varepsilon_t d_t(\hat{\beta}) + o_p(1) \\ &= \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) H_t^{i,l} \left(\hat{\beta}_l - \phi_l^0 \right) + \sum_{i=1}^4 \sum_{t=1}^T 2 \left(\hat{\beta}_i - \beta_i^0 \right) \varepsilon_t h_t^{(i)} + o_p(1). \end{aligned}$$

Ahora, es directo probar que $h_t^{(i)}$ son secuencias $L_2 - mixingala$ de tamaño $-1/2$ con constantes finitas. Por lo tanto, siguiendo los mismos pasos que en el Teorema 2.3 para el caso Observable, se obtiene que

$$\frac{1}{T} \sum_{t=1}^T H_t^{i,l} \xrightarrow{p} H^{i,l}(\beta^0) = E\left(H_t^{i,l}\right),$$

$$D_T(\beta_0) = \frac{1}{T^{1/2}} \sum_{t=1}^T \varepsilon_t h_t^{(i)} \xrightarrow{d} N\left(\vec{0}, \Omega\right),$$

con $\Omega(i, l) = \sigma_\varepsilon^2 E\left(h_t^{(i)} h_t^{(l)}\right)$. Además, dado que las proposiciones 3.11 y 3.12 son ciertas para todo $\tilde{\beta}$ tal que $T^{1/2}(\tilde{\beta} - \beta^0) = O_p(1)$, podemos definir $T^{1/2}(\tilde{\beta} - \beta^0) = -H_T^{-1}(\beta^0) D_T(\beta^0)$, tal que

$$Q_T(\tilde{\beta}) - Q_T(\beta^0) = T(\tilde{\beta} - \beta^0)' H(\beta^0)(\tilde{\beta} - \beta^0) + 2T^{1/2}(\tilde{\beta} - \beta^0)' D_T(\beta^0) + o_p(1).$$

Siguiendo la argumentación del Teorema 2.3, concluimos que

$$T^{1/2}(\hat{\beta} - \beta^0) \xrightarrow{d} N\left(\vec{0}, \sigma_\varepsilon^2 H^{-1}(\beta^0)\right).$$

■

Prueba del Teorema 3.4. Como en el caso Observable, la prueba de $\hat{\sigma}_\varepsilon^2 - \sigma_\varepsilon^2 = o_p(1)$ es directa. Por lo tanto, debemos probar que para $\forall \eta > 0$, $\exists \nu_T$ tal que

$$\lim_{T \rightarrow \infty} P\left[\varrho\left(\beta_0, \hat{\beta}\right) < \nu_T\right] = 1 \text{ y}$$

$$\lim_{T \rightarrow \infty} P\left[\sup_{\beta / \varrho(\beta_0, \hat{\beta}) < \nu_T} T^{-1} \left| \sum_{t=1}^T h_t^{(i)}(\beta) h_t^{(j)}(\beta) - h_t^{(i)}(\beta^0) h_t^{(j)}(\beta^0) \right| > \eta\right] = 0,$$

con

$$h_t^{(1)}(\beta) = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k} e_{t-1-j}^+(\beta), \quad h_t^{(2)}(\beta) = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k} e_{t-1-j}^-(\beta),$$

$$h_t^{(3)}(\beta) = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k} x_{t-1-j}, \quad h_t^{(4)}(\beta) = \sum_{j=0}^t \prod_{k=1}^j \theta_{t-k}.$$

Como es fácilmente comprobable,

$$\sup_{\beta/\varrho(\beta_0, \beta) < \nu_T} \left| h_t^{(i)}(\beta) - h_t^{(i)}(\beta_0) \right| \leq \bar{s}_t(\beta_0, \nu_T)$$

$$\sup_{\beta/\varrho(\beta_0, \beta) < \nu_T} \left| h_t^{(i)}(\beta) \right| \leq \bar{h}_t,$$

con

$$\begin{aligned} \bar{s}_t(\beta_0, \nu_T) &= 2 \sum_{j=1}^t \prod_{k=1}^j \theta_{t-k}^0 \bar{v}_{t-j}(\beta_0, \nu_T) + \\ &+ \sum_{j=1}^t \prod_{k=1}^j \theta_{t-k}^0 [\nu + 1 (\bar{A}_{3,t-j}(\beta_0, \nu_T)) + 1 (\bar{A}_{4,t-j}(\beta_0, \nu_T))] (1 + |x_{t-j}| + |\varepsilon_{t-j}| + \bar{h}_{t-j}) \\ \bar{h}_t &= 1 + |x_{t-j}| + \bar{e}_{t-1} + \lambda_1 \bar{h}_{t-1}. \end{aligned}$$

Por la desigualdad de Minkowsky, la LIE y los supuestos del Teorema,

$$\|\bar{s}_t(\beta_0, \nu_T)\|_2 \leq \nu_T^{1/2} K_s$$

$$\|\bar{h}_t\|_2 < \infty.$$

Por el Teorema 3.2, sabemos que $T^{1/2}(\hat{\beta}_i - \beta_i^0) = O_p(1)$, con lo que existirá $0 < d < 1/2$ tal que $\lim_{T \rightarrow \infty} P[\varrho(\hat{\beta}, \beta_0) < \nu_T] = 1$ para $\nu_T = T^d$. Ahora, para todo $\eta > 0$ tenemos

$$P \left[\sup_{\beta/\varrho(\beta_0, \beta) < \nu} \frac{1}{T} \left| \sum_{t=1}^T h_t^{(i)}(\beta) h_t^{(j)}(\beta) - h_t^{(i)}(\beta_0) h_t^{(j)}(\beta_0) \right| > \eta \right] \leq P \left[\frac{1}{T} \left| \sum_{t=1}^T \bar{s}_t(\beta_0, \nu_T) \bar{h}_t \right| > \eta/2 \right].$$

por lo tanto,

$$P \left[\frac{1}{T} \left| \sum_{t=1}^T \bar{s}_t(\beta_0, \nu_T) \bar{h}_t \right| > \eta/2 \right] \leq \frac{2 \|\bar{s}_t(\beta_0, \nu_T)\|_2 \|\bar{h}_t\|_2}{\eta} \leq \frac{2 \nu_T^{1/2} K_s \|\bar{h}_t\|_2}{\eta} \rightarrow 0.$$

Con este resultado, y dado que en el Teorema 3.3 se prueba la convergencia de

$\frac{1}{T} \sum_{t=1}^T h_t^{(i)}(\beta_0) h_t^{(j)}(\beta_0)$, queda probado este Teorema. ■

Inferencia cuando r^0 es desconocido

Asumiendo que $|r^0| > 0$ aunque desconocido, se propone realizar la siguiente regresión:

$$x_t = \alpha_1 + \alpha_2 x_{t-1} + \alpha_3 \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) > \hat{r}) + \alpha_4 \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) < \hat{r}) + u_t,$$

que en términos matriciales se puede expresar como,

$$X_T = Z_T \alpha + U_T \quad (3.25)$$

donde $\alpha' = (\alpha_1, \dots, \alpha_4)$, $X_T' = (x_2, \dots, x_T)$, $U_T' = (u_1, \dots, u_{T-1})$ y $Z_T' = (Z_{T,1}, \dots, Z_{T,T-1})$ con $Z_{T,t-1}' = (1, x_{t-1}, \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) > \hat{r}), \hat{\varepsilon}_{t-1} 1(g(\hat{\varepsilon}_{t-1}) < \hat{r}))$ y $\hat{\varepsilon}_t$ el estimador CLS de la primera etapa. Definiendo $\alpha^{0'} = (\mu^0, \rho^0, \theta_1^0, \theta_2^0)$, $U_T^{0'} = (\varepsilon_{T-1}, \dots, \varepsilon_1)$ y $Z_T^{0'} = (Z_{T,1}^0, \dots, Z_{T,T-1}^0)$, con $Z_{T,t-1}^{0'} = (1, x_{t-1}, \varepsilon_{t-1} 1(g(\varepsilon_{t-1}) > r^0), \varepsilon_{t-1} 1(g(\varepsilon_{t-1}) < r^0))$, X_T puede escribirse como,

$$X_T = Z_T^0 \alpha^0 + U_T^0.$$

Usando que la matriz $E[Z_t^0 Z_t^{0'}]$ es definida positiva, lo cual se deriva de los supuestos ASZ.2 y ASG.2, que Z_t^0 es estacionaria, ε_t es *iid*, y con el resto de supuestos del Teorema 3.5, es fácil probar que

$$\begin{aligned} \tilde{\alpha} &= (Z_T^{0'} Z_T^0)^{-1} Z_T^{0'} X_T \xrightarrow{p} \alpha^0, \\ T^{-1/2} (\tilde{\alpha} - \alpha^0) &\xrightarrow{d} N\left(0, \sigma_\varepsilon^2 E(Z_t^{0'} Z_t^0)^{-1}\right), \end{aligned}$$

donde $\sigma_\varepsilon^2 = E(\varepsilon_t^2)$.

Prueba del Teorema 3.5. Dado el resultado anterior, el Teorema será cierto si

$T^{-1/2}(\tilde{\alpha} - \hat{\alpha}) \xrightarrow{p} 0$. Para probar esto, definimos

$$Z'_T(\phi) = (Z_{T,1}(\phi), \dots, Z_{T,T-1}(\phi))$$

$$Z'_{T,t-1}(\phi) = (1, x_{t-1}, e_{t-1}(\phi) 1(g(e_{t-1}(\phi)) > \phi_5), e_{t-1}(\phi) 1(g(e_{t-1}(\phi)) < \phi_5))$$

$$V_T(\phi, \phi^0) = (0, 0, V_{1,T}(\phi, \phi^0), V_{2,T}(\phi, \phi^0))$$

$$V_{1,T,t-1}(\phi, \phi^0) = \varepsilon_{t-1} [1(g(e_{t-1}(\phi)) > \phi_5) - 1(g(\varepsilon_{t-1}) > r^0)] + \\ + v_{t-1}(\phi, \phi^0) 1(g(e_{t-1}(\phi)) > \phi_5)$$

$$V_{2,T,t-1}(\phi, \phi^0) = \varepsilon_{t-1} [1(g(e_{t-1}(\phi)) < \phi_5) - 1(g(\varepsilon_{t-1}) < r^0)] + \\ + v_{t-1}(\phi, \phi^0) 1(g(e_{t-1}(\phi)) < \phi_5)$$

donde $v_{t-1}(\phi, \phi^0)$ esta definido en el Teorema 3.1, para el caso del tamaño, y en el Teorema 3.2, para el caso del signo. En primer lugar probaremos que

$$T^{-1} [(Z'_T Z_T) - (Z_T^{0'} Z_T^0)] = T^{-1} \left[\left(Z'_T(\hat{\phi}) Z_T(\hat{\phi}) \right) - (Z_T^{0'} Z_T^0) \right] \xrightarrow{p} 0.$$

Por las definiciones anteriores,

$$\left[\left(Z'_T(\hat{\phi}) Z_T(\hat{\phi}) \right) - (Z_T^{0'} Z_T^0) \right] = V'_T(\hat{\phi}, \phi^0) V_T(\hat{\phi}, \phi^0) + Z'_{T,t-1}(\phi^0) V_T(\hat{\phi}, \phi^0) + \\ + V'_T(\hat{\phi}, \phi^0) Z_{T,t-1}(\phi^0).$$

Como en casos anteriores, definimos los siguientes procesos cota para $V_T(\phi, \phi^0)$, $V_{1,T}(\phi, \phi^0)$ y $V_{2,T}(\phi, \phi^0)$,

$$\bar{V}_{1,T,t-1}(\phi^0, \nu) = \bar{V}_{2,T,t-1}(\phi^0, \nu) = |\varepsilon_{t-1}| [1(\bar{A}_{3,t-1}(\phi^0, \nu)) + 1(\bar{A}_{4,t-1}(\phi^0, \nu))] + \bar{v}_{t-1}(\phi^0, \nu) \\ \bar{V}_T(\phi^0, \nu) = (0, 0, \bar{V}_{1,T}(\phi^0, \nu), \bar{V}_{2,T}(\phi^0, \nu)),$$

donde $[1(\bar{A}_{3,t-1}(\phi^0, \nu)) + 1(\bar{A}_{4,t-1}(\phi^0, \nu))]$ esta definido en la ecuación (3.10), para el caso del tamaño, y (3.20) para el caso del signo, de forma que

$$\sup_{\phi/\varrho(\phi, \phi^0) < \nu_T} [(Z'_T(\phi) Z_T(\phi)) - (Z_T^{0'} Z_T^0)] \leq \bar{V}'_T(\phi^0, \nu_T) \bar{V}_T(\phi^0, \nu_T) + |Z'_{T,t-1}(\phi^0)| \bar{V}_T(\phi^0, \nu_T) + \\ + \bar{V}'_T(\phi^0, \nu_T) |Z_{T,t-1}(\phi^0)|.$$

Siguiendo la demostración de los lemas 3.3, para el caso del tamaño, y 3.6, para el caso del signo, es fácil probar que para $i = 3, 4$ y $k, k' \leq 4$,

$$\left\| \varepsilon_{t-1}^{k'} 1 \left(\bar{A}_{i,t-1} (\phi^0, \nu_T) \right) \right\|_k \leq v_T^{1/k} K_A,$$

por lo tanto, tomado $v_T = T^{-d}$ con $d \in (0, 1)$,

$$\begin{aligned} \left\| \bar{V}'_T (\phi^0, v_T) \bar{V}_T (\phi^0, v_T) \right\|_1 &= O(T^{-d/2}) \\ \left\| Z'_{T,t-1} (\phi^0) \bar{V}_T (\phi^0, v_T) \right\|_1 &= O(T^{-d/2}). \end{aligned}$$

entendiendo que $\|X\|_k = O(T^{-d/2})$, con X una matriz $n \times m$, si la igualdad se cumple para cada uno de sus elementos. Con estos dos resultados, es directo probar que

$$\sup_{\phi/\varrho(\phi, \phi^0) < v_T} T^{-1} \left[(Z'_T(\phi) Z_T(\phi)) - (Z_T^{0'} Z_T^0) \right] \xrightarrow{p} 0,$$

y dado que $P \left[\varrho(\hat{\phi}, \phi^0) < v_T \right] \xrightarrow{T \rightarrow \infty} 1$, por los Teoremas 3.1 y 3.2, se obtiene que

$$T^{-1} \left[\left(Z'_T(\hat{\phi}) Z_T(\hat{\phi}) \right) - (Z_T^{0'} Z_T^0) \right] \xrightarrow{p} 0.$$

Ahora probamos que

$$T^{-1/2} \left[Z'_T X_T - Z_T^{0'} X_T \right] \xrightarrow{p} 0.$$

Para ello escribimos el proceso como,

$$\begin{aligned} \sup_{\phi/\varrho(\phi, \phi^0) < v_T} T^{-1/2} \left[(Z_T^{0'} + V'_T(\phi, \phi^0)) X_T - Z_T^{0'} X_T \right] &\leq T^{-1/2} \bar{V}'_T(\phi^0, v_T) |Z_T^0 \beta^0| + \\ &\quad + \bar{V}'_T(\phi^0, v_T) |U_T^0|, \end{aligned}$$

con $v_T = T^{-d}$ y $d \in (3/4, 1)$. Usando la desigualdad de Hölder's y los resultados anteriores,

para $i = 1, 2$ y $\forall \eta > 0$,

$$P \left[\left| \frac{\sum_{t=1}^T \bar{V}'_{i,T,t-1}(\phi^0, v_T) |\varepsilon_{t-1}|}{T^{1/2}} \right| > \eta \right] \leq T^{1/2} \left\| \bar{V}'_{1,T,t-1}(\phi^0, v_T) \right\|_{3/2} \|\varepsilon_{t-1}\|_3 = O\left(T^{\frac{1}{2} - \frac{2}{3}d}\right)$$

$$P \left[\left| \frac{\sum_{t=1}^T \bar{V}'_{i,T,t-1}(\phi^0, v_T) |\varepsilon_t|}{T^{1/2}} \right| > \eta \right] \leq T^{1/2} \left\| \bar{V}'_{1,T,t-1}(\phi^0, v_T) \right\|_{3/2} \|\varepsilon_t\|_3 = O\left(T^{\frac{1}{2} - \frac{2}{3}d}\right).$$

Ahora, para $d > 3/4$,

$$T^{-1/2} \left[\bar{V}'_T(\phi^0, \nu_T) |Z_T^0 \beta^0| + \bar{V}'_T(\phi^0, \nu_T) |U_T^0| \right] \xrightarrow{p} 0$$

$$\sup_{\phi/\varrho(\phi, \phi^0) < v_T} T^{-1/2} \left[(Z'_{T,t-1}(\phi^0) + V'_T(\phi, \phi^0)) X_T - Z_T^{0'} X_T \right] \xrightarrow{p} 0,$$

y dado que $\left[\varrho(\hat{\phi}, \phi^0) < v_T \right] \xrightarrow{T \rightarrow \infty} 1$,

$$T^{-1/2} \left[Z'_T X_T - Z_T^{0'} X_T \right] \xrightarrow{p} 0.$$

Por lo tanto, usando el Teorema de Slutsky,

$$T^{1/2} \left(\hat{\beta} - \beta^0 \right) \xrightarrow{d} N \left(0, \sigma_\varepsilon^2 E \left(Z_t^{0'} Z_t^0 \right)^{-1} \right).$$

Con esto, queda probado el Teorema. ■

Contraste de existencia de Threshold para $|r^0| > 0$ desconocido.

En la prueba nos centramos en el caso del tamaño. El caso del signo se obtiene de forma similar. Como en el caso Observable, primero probaremos que el $\sup_{r \in [\underline{r}, \bar{r}]} \sqrt{T} \left(\varrho \left(\hat{\beta}(r), \beta^0 \right) \right) = O_p(1)$, con $\beta^0 = (\theta, \alpha, \rho, \mu)$, ya que bajo la nula de no efecto threshold los estimadores de ϕ^0 no serán T consistentes. Durante la prueba, asumiremos que $\sup_{r \in [\underline{r}, \bar{r}]} \left(\varrho \left(\hat{\beta}(r), \beta^0 \right) \right) = o_p(1)$, lo cual se prueba fácilmente siguiendo la primera parte del Teorema 3.1. Por lo tanto, podemos escoger $\epsilon > 0$ suficientemente pequeño tal que para T suficientemente grande,

$P \left[\sup_{r \in [\underline{r}, \bar{r}]} \left(\varrho \left(\widehat{\beta}(r), \beta^0 \right) \right) \leq \epsilon \right] \geq 1 - \epsilon$. Igualmente, el resultado se obtendrá de las proposiciones 2.19 y 2.20, usadas en el caso Observable. Antes de probar ambas proposiciones para este caso, estudiamos los procesos que usaremos como cotas. Definimos $\beta = (\theta, \alpha, \rho, \mu)$, y para un r dado,

$$\begin{aligned}
 e_t(\beta, r) &= x_t - \rho x_{t-1} - \mu + \theta e_{t-1}(\beta, r) + \alpha e_{t-1}(\beta, r) 1(B_{t-1}) \\
 e_t(\beta, r) - \varepsilon_t &= v_t(\beta, \beta^0, r) = (\rho^0 - \rho) x_{t-1} + (\mu^0 - \mu) + (\theta - \theta^0) e_{t-1}(\beta, r) + \\
 &\quad + \alpha e_{t-1}(\beta, r) 1(B_{t-1}) + \theta^0 v_{t-1}(\beta, \beta^0, r) \\
 &= \sum_{j=1}^t \theta_0^{j-1} [(\rho^0 - \rho) x_{t-j} + (\mu^0 - \mu) + (\theta - \theta^0) e_{t-j}(\beta, r) + \alpha e_{t-j}(\beta, r) 1(B_{t-j})] \\
 &= \sum_{i=1}^4 (\beta_i - \beta_0^i) h_t^{(i)}(\beta, r)
 \end{aligned} \tag{3.26}$$

con $B_{t-j} = 1(|e_{t-j}(\beta, r)| < r)$. Por lo tanto,

$$\begin{aligned}
 \sup_{\beta} \sup_{r \in [\underline{r}, \bar{r}]} e_t(\beta, r) &\leq \bar{e}_t = |x_t| + |x_{t-1}| + \bar{\mu} + \alpha \bar{r} + (1 - \delta) \bar{e}_{t-1} \\
 \sup_{\beta / \varrho(\beta, \beta^0) < \nu} \sup_{r \in [\underline{r}, \bar{r}]} v_t(\beta, \beta^0, r) &\leq \bar{v}_t(\beta^0, r, \nu) = \nu \bar{h}_t(r) \\
 \bar{h}_t &= \sum_{j=1}^t \theta_0^{j-1} [|x_t| + 1 + 2\bar{e}_{t-j}] \\
 s_t^{(i)}(\phi, \phi^*) &= \sum_{i=1}^4 h_t^{(i)}(\phi) - h_t^{(i)}(\phi^*) = \sum_{j=1}^t \theta_0^{j-1} v_{t-j}(\phi, \phi^*) + \\
 &\quad + \sum_{j=1}^t \theta_0^{j-1} [1(A_{3,t-j}(\phi, \phi^*)) e_{t-j}^* + 1(A_{4,t-j}(\phi, \phi^*)) e_{t-j}] \\
 \sup_{\phi / \varrho(\phi, \phi^*) < \nu} s_t^{(i)}(\phi, \phi^*) &\leq \bar{s}_t(\phi^*, \nu) = \sum_{j=1}^t \theta_0^{j-1} \left[\bar{v}_{t-j}(\phi^*, \nu) + \bar{r}^* \left[\sum_{i=3}^4 1(\bar{A}_{i,t-j}(\phi^*, \nu)) \right] \right],
 \end{aligned}$$

donde $\bar{v}_{t-j}(\phi^*, \nu)$ y $[1(\bar{A}_{3,t-j}(\phi^*, \nu)) + 1(\bar{A}_{4,t-j}(\phi^*, \nu))]$ fueron definidas en el Teorema 3.1. Como entonces, $e_t(\beta, r)$, \bar{e}_t y $\bar{v}_{t-j}(\phi^*, \nu)$ serán $L_2 - NED$, con el mismo tamaño y

constantes. Respecto a los momentos,

$$\begin{aligned}
 E(\bar{e}_t(r))^2 &\leq \frac{2E|x_t|^2 + \bar{\mu} + \bar{r}}{\delta^2} \leq K_e \\
 E(\bar{h}_t(r))^2 &\leq \frac{2K_e}{(1 - \theta_0)^2} \leq K_h \\
 E(\bar{v}_t(\beta^0, r, \nu))^2 &\leq \nu^2 K_h \\
 E(\bar{v}_t(\phi^*, \nu) | \varepsilon_{t-1}|)^2 &\leq \nu K_{v\varepsilon} \Rightarrow E(\bar{v}_t(\phi^*, \nu) | e_{t-1}(\beta^*, r)|)^2 \leq \nu K_{ve} \\
 E(\bar{s}_t(\phi^*, \nu))^2 &\leq \nu K_s.
 \end{aligned}$$

Siguiendo los mismos pasos que en el Teorema 3.1, se puede probar que $\bar{v}_t(\beta^0, r, \nu)$ y $\bar{s}_t(\phi^*, \nu)$ son L_2 -mixingales de tamaño $-1/2$ con constantes proporcionales a ν y $\nu^{1/2}$, respectivamente. Finalmente, definimos $B_{\Delta, \epsilon} = \left\{ \phi \in \Theta : \left(\frac{\Delta}{T} \right)^{1/2} < \varrho(\beta, \beta^0) < \epsilon, \quad r \in [\underline{r}, \bar{r}] \right\}$.

Las proposiciones necesarias son,

Proposición 3.13 *Para todo $\eta, \varepsilon > 0$, $\exists \Delta < \infty$ tal que*

$$P \left[\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t(v_t(\beta, \beta^0, r))}{T\varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon$$

Proposición 3.14 *Para todo $\varepsilon > 0$, $\exists \Delta < \infty$ y $\eta > 0$ tal que*

$$P \left[\inf_{\phi \in B_{\Delta, \epsilon}} \frac{\sum_{t=1}^T (v_t(\beta, \beta^0, r))^2}{T\varrho^2(\beta, \beta^0)} > \eta \right] \geq 1 - \varepsilon$$

Prueba de la Proposición 3.13. Usando los últimos resultados,

$$\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t(v_t(\beta, \beta^0, r))}{T\varrho^2(\beta, \beta^0)} \right| \leq \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \alpha 1(B_{t-j}) \varepsilon_{t-j}}{T\varrho^2(\beta, \beta^0)} \right| + \quad (3.27)$$

$$+ \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} [(\theta - \theta^0) + \alpha 1(B_{t-j})] v_{t-j}(\beta, \beta^0, r)}{T\varrho^2(\beta, \beta^0)} \right| + \quad (3.28)$$

$$+ \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} (\rho^0 - \rho) x_{t-j} + (\mu^0 - \mu) + (\theta - \theta^0) \varepsilon_{t-j}}{T\varrho^2(\beta, \beta^0)} \right| \quad (3.29)$$

Ahora, probaremos que cada uno de los supremos de las ecuaciones (3.27) a (3.29) cumple la proposición 3.13. Usando la partición, $B_{\epsilon, \ell_j}^\Delta(\phi)$, construida a partir de $B_{\epsilon, j_l}^l(\Delta, \phi_l) =$

$\left\{ \phi_l \in B_{\Delta, \epsilon}^+ \middle/ 0 \leq \frac{\phi_l - \phi_l^{j_l}}{\phi_l^{j_l} - \phi_l^0} \leq (b-1) \right\}$ con $\phi_l^{j_l} = \phi_l^0 + \frac{\Delta}{T^{1/2}} b^{j_l}$ para $l \leq 4$, $B_{\epsilon, j_5}^5(\Delta, \phi_5) = \left\{ \phi_5 \in [\underline{r}, \bar{r}] \middle/ 0 \leq \phi_5 - \phi_5^{j_5} \leq T^{-1/2} \Delta (b-1) \right\}$ y $\phi_5^{j_5} = \underline{r} + \frac{\Delta}{T^{1/2}} (b-1) j_5$, la ecuación (3.27) puede acotarse por,

$$\begin{aligned} & \sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \alpha 1(B_{t-j}) \varepsilon_{t-j}}{T \varrho^2(\beta, \beta^0)} \right| \leq \\ & 4 \max_{\ell_j} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} 1(|e_{t-1}(\phi^{\ell_j})| < r^{\ell_j}) \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} \right| + \\ & + 4 \max_{\ell_j} \sup_{\phi \in B_{\epsilon, \ell_j}^{\Delta}(\phi)} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} [1(|e_{t-1}(\phi^{\ell_j})| < r^{\ell_j}) - 1(|e_{t-1}(\phi)| < r)] \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} \right| \end{aligned} \quad (3.30)$$

$$(3.31)$$

para b suficientemente cercano a 1. Para el máximo de la ecuación (3.30), tenemos que,

$$\begin{aligned} & \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} 1(|e_{t-1}(\phi^{\ell_j})| < r^{\ell_j}) \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} = \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r^{\ell_j}) \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} \\ & + \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} [1(|e_{t-1}(\phi^{\ell_j})| < r^{\ell_j}) - 1(|\varepsilon_{t-j}| < r^{\ell_j})] \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)}. \end{aligned} \quad (3.32)$$

Ahora, usando

$$\| [1(|e_{t-j}(\phi^{\ell_j})| < r^{\ell_j}) - 1(|\varepsilon_{t-j}| < r^{\ell_j})] \varepsilon_t^2 \varepsilon_{t-j}^2 \|_2 \leq \varrho^{1/2}(\beta^{\ell_j}, \beta^0) K_e,$$

se tiene que,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{m-1} \theta^{0j-1} [1(|e_{t-1}(\phi^{\ell_j})| < r^{\ell_j}) - 1(|\varepsilon_{t-j}| < r^{\ell_j})] \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} \right| > \eta \right] & \leq \\ & \leq \sum_{j=1}^{\infty} \frac{j^3 \bar{K}_e}{\Delta b^j \eta^2} \end{aligned}$$

que será tan pequeño como queramos para Δ suficientemente grande. Para el primer sumando de la ecuación (3.32),

$$\max_{\ell_j} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=m}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r^{\ell_j}) \varepsilon_{t-j}}{T \varrho(\beta^{\ell_j}, \beta^0)} \right| \leq \max_{j_5} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=m}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r^{j_5}) \varepsilon_{t-j}}{T^{1/2} \Delta b} \right|.$$

Asumiendo que $T^{1/2} \sum_{t=1}^T 2\varepsilon_t \sum_{j=m}^t \theta^{0j-1} 1(|\varepsilon_{t-j}| < r) \varepsilon_{t-j}$ es tight en r , tenemos que para todo $\eta > 0$,

$$P \left[\max_{j_5} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=m}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r^{j_5}) \varepsilon_{t-j}}{T^{1/2}} \right| > \eta \right] \leq \varepsilon^*,$$

y por lo tanto, para Δ suficientemente grande

$$P \left[\max_{j_5} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=m}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r^{j_5}) \varepsilon_{t-j}}{T^{1/2} \Delta b} \right| > \eta \right] \leq \varepsilon,$$

con $\varepsilon > 0$ tan pequeño como queramos. Como en el caso Observable, la tightness de $T^{1/2} \sum_{t=1}^T 2\varepsilon_t \sum_{j=m}^{t-1} \theta^{0j-1} 1(|\varepsilon_{t-j}| < r) \varepsilon_{t-j}$ se probará más adelante.

Para el máximo de la ecuación (3.31) definimos,

$$\Psi_t(\phi^{\ell_j}) = |\varepsilon_t| \sum_{j=1}^{t-1} \theta^{0j-1} |\varepsilon_{t-j}| [1(\bar{A}_{3,t-j}(\phi^{\ell_j}, \nu(\phi^{\ell_j}))) + 1(\bar{A}_{4,t-j}(\phi^{\ell_j}, \nu(\phi^{\ell_j})))]$$

$$\Psi_t(\phi^{\ell_j}) \geq \sup_{\phi \in B_{\varepsilon, \ell_j}^{\Delta}(\phi)} \varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \varepsilon_{t-j} [1(|e_{t-j}(\phi^{\ell_j})| < r^{\ell_j}) - 1(|e_{t-j}(\phi)| < r)]$$

$$E(\Psi_t(\phi^{\ell_j})) \leq \nu(\phi^{\ell_j}) K_{\Psi}$$

$$E(\Psi_t^2(\phi^{\ell_j})) \leq \nu(\phi^{\ell_j}) K_{\Psi}$$

donde $\nu(\phi^{\ell_j}) \leq \rho(\phi^{\ell_j}, \phi) \leq \rho(\beta^{\ell_j}, \beta^0) (b-1)$ y $\Psi_t(\phi^{\ell_j})$ es L_2 -mixingala de tamaño $-1/2$ con constante $T^{-1/4} K_{\Psi}^{1/2}$ lo cual se puede probar de forma similar a $\bar{v}_t(\phi^*, \nu)$ en el Teorema 3.1. Por lo tanto,

$$P \left[\max_{\ell_j} \frac{\sum_{t=1}^T [\Psi_t(\phi^{\ell_j}) - E(\Psi_t(\phi^{\ell_j}))]}{T \varrho(\beta^{\ell_j}, \beta^0)} > \eta \right] \leq \sum_{k=1}^{\frac{T^{1/2}(\bar{r}-r)}{\Delta}} \sum_{j=1}^{\infty} \frac{T^{-1/2} j^3 \bar{K}_{\Psi}}{\Delta b^j \eta^2} \leq \sum_{j=1}^{\infty} \frac{j^3 \bar{K}_{\Psi}}{\Delta^2 b^j \eta^2}$$

$$\max_{\ell_j} \frac{\sum_{t=1}^T E(\Psi_t(\phi^{\ell_j}))}{T \varrho(\beta^{\ell_j}, \beta^0)} \leq K_{\Psi} (b-1).$$

Para Δ suficientemente grande, la primera probabilidad será tan pequeña como queramos. Además, para b suficientemente cercano a 1, el máximo de la segunda desigualdad

también será tan pequeño como queramos. Resumiendo, para el supremo de (3.27) se cumple que para $\forall \eta, \varepsilon > 0 \exists \Delta$ suficientemente grande tal que

$$P \left[\sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \sum_{j=1}^{t-1} \theta^{0j-1} \alpha 1(B_{t-j}) \varepsilon_{t-j}}{T \varrho^2(\beta, \beta^0)} \right| > \eta \right] \leq \varepsilon.$$

Para el supremo de (3.28),

$$\begin{aligned} & \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} (\theta_1 - \theta_1^0) v_{t-j}(\beta, \beta^0, r)}{T \varrho^2(\beta, \beta^0)} \right| \\ & \leq \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \sum_{i=1}^4 h_{t-j}^{(i)}(\beta, r)}{T} \right|, \end{aligned}$$

y

$$\begin{aligned} & \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \alpha 1(B_{t-j}) v_{t-j}(\beta, \beta^0, r)}{T \varrho^2(\beta, \beta^0)} \right| \\ & \leq \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \sum_{i=1}^4 1(B_{t-j}) h_{t-j}^{(i)}(\beta, r)}{T} \right| \end{aligned}$$

En la primera desigualdad, para $h_t^{(i)}(\beta, r)$ con $i = 3, 4$, la prueba de que el supremo convergerá a 0 es directa, dado que $h_t^{(i)}(\beta, r)$ no dependerá de r y β . Para $i = 1, 2$,

$$\begin{aligned} & \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} h_{t-j}^{(1)}(\beta, r)}{T \varrho(\beta, \beta^0)} \right| \leq \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \sum_{k=1}^{t-j-1} \theta^{0k-1} e_{t-k}(\beta, r)}{T} \right| \\ & \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} h_{t-j}^{(2)}(\beta, r)}{T \varrho(\beta, \beta^0)} \right| \\ & \leq \sup_{\phi \in B_{\Delta, \varepsilon}} \left| \frac{\sum_{t=1}^T \varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} \sum_{k=1}^{t-j-1} \theta^{0k-1} 1(B_{t-k}) e_{t-k}(\beta, r)}{T} \right|. \end{aligned}$$

Ahora, dado (β, r) , ambos supremos convergerán a 0, por lo tanto, usando las propiedades de $e_{t-k}(\beta, r)$ descritas en el Teorema 3.1, podemos obtener la ULLN, que garantiza la convergencia del supremo. La prueba será similar a los casos anteriores.

Finalmente, para el supremo de (3.29),

$$\sup_{\phi \in B_{\Delta, \epsilon}} \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} (\rho^0 - \rho) \theta^{0j-1} x_{t-j}}{T \varrho^2(\beta, \beta^0)} \right| \leq \left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} x_{t-j}}{T^{1/2} \Delta} \right|$$

$$P \left[\left| \frac{\sum_{t=1}^T 2\varepsilon_t \sum_{j=1}^{t-1} \theta^{0j-1} x_{t-j}}{T^{1/2} \Delta} \right| > \eta \right] \leq \frac{K_{\varepsilon e}}{\Delta^2 \eta^2},$$

que será tan pequeño como queramos para Δ suficientemente grande. El mismo resultado se obtiene para los procesos $\varepsilon_t \sum_{j=1}^{t-1} (\mu^0 - \mu) \theta^{0j-1}$ y $\varepsilon_t \sum_{j=1}^{t-1} (\theta^0 - \theta) \theta^{0j-1} \varepsilon_{t-j}$.

Con los resultados para los términos (3.27) a (3.29) queda probada la proposición 3.13.

■

Prueba de la Proposición 3.14. La prueba de esta proposición se obtiene de forma paralela a la de la proposición 2.20 en el Teorema 2.6, y por motivo de espacio es omitida.

■

De ambas proposiciones, 3.13 y 3.14, se tiene que

$$\sup_{r \in [\underline{r}, \bar{r}]} \sqrt{T} \varrho_1 \left(\hat{\beta}(r), \beta^0 \right) = O_p(1).$$

Para obtener la distribución asintótica propuesta en el Teorema, expresamos $Q_T(\hat{\phi}, r) - Q_T(\phi^0)$ como,

$$\begin{aligned} Q_T(\hat{\beta}, r) - Q_T(\beta^0) &= \sum_{t=1}^T v_{t-1}^2 \left(\hat{\beta}, \beta^0, r \right) + 2 \sum_{t=1}^T \varepsilon_t v_{t-1} \left(\hat{\beta}, \beta^0, r \right) \\ &= \sum_{t=1}^T \left(\sum_{i=1}^4 \left(\hat{\beta}_i - \beta_i^0 \right) h_t^{(i)} \left(\hat{\beta}, r \right) \right)^2 + 2 \sum_{t=1}^T \varepsilon_t \sum_{i=1}^4 \left(\hat{\beta}_i - \beta_i^0 \right) h_t^{(i)} \left(\hat{\beta}, r \right) \\ &= \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T \left(\hat{\beta}_i - \beta_i^0 \right) H_{t-1}^{i,l} \left(\hat{\beta}, r \right) \left(\hat{\beta}_l - \beta_l^0 \right) + \\ &\quad + 2 \sum_{t=1}^T \varepsilon_t \sum_{i=1}^4 \left(\hat{\beta}_i - \beta_i^0 \right) h_t^{(i)} \left(\hat{\beta}, r \right) \end{aligned}$$

Ahora, como en el caso Observable, necesitamos el siguiente resultado

$$\begin{aligned} Q_T(\widehat{\beta}, r) - Q_T(\beta^0) &= \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\widehat{\beta}_i - \beta_i^0) H_{t-1}^{i,l}(\beta^0, r) (\widehat{\beta}_l - \beta_l^0) + \\ &+ 2 \sum_{t=1}^T \varepsilon_t \sum_{i=1}^4 (\widehat{\beta}_i - \beta_i^0) h_t^{(i)}(\beta^0, r) + o_p(1), \end{aligned}$$

con el $o_p(1)$ uniformemente en r . Para ello, usaremos los siguientes dos lemas.

Lema 3.10 *Bajo los supuestos del Teorema 3.6,*

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \left(h_t^{(i)}(\widehat{\beta}(r)) - h_t^{(i)}(\beta^0, r) \right) \right| = o_p(1)$$

para $i = 1, \dots, 4$.

Prueba del Lema 3.10. El lema se cumplirá si para $\forall \eta > 0$,

$$\lim_{T \rightarrow \infty} P \left[\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \left(h_t^{(i)}(\widehat{\beta}(r)) - h_t^{(i)}(\beta^0, r) \right) \right| > \eta \right] = 0.$$

Para $i = 3, 4$, la prueba es directa dado que $h_t^{(i)}(\widehat{\beta}(r))$ no depende de (β, r) y $(h_t^{(i)}(\widehat{\beta}(r)) - h_t^{(i)}(\beta^0, r)) = 0$. Para $i = 1$,

$$\begin{aligned} \sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \left(h_t^{(1)}(\widehat{\beta}(r)) - h_t^{(1)}(\beta^0, r) \right) \right| &= \sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} v_{t-j}(\widehat{\beta}, \beta^0, r) \right| \\ &= \sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} \sum_{i=1}^4 (\widehat{\beta}_i - \beta_i^0) h_{t-j}^{(i)}(\widehat{\beta}(r)) \right|. \end{aligned}$$

Ahora, por los resultados previos, $\sup_{r \in [\underline{r}, \bar{r}]} \sqrt{T} \varrho(\widehat{\beta}(r), \beta^0) = O_p(1)$, por lo tanto, será suficiente probar que para $i = 1, \dots, 4$,

$$\lim_{T \rightarrow \infty} P \left[\sup_{\phi \in B_\Delta} \frac{\Delta}{T} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0^{j-1}} h_{t-j}^{(i)}(\beta, r) \right| > \eta \right] = 0,$$

para $B_\Delta = \{\phi \in \Theta / |\beta_l - \beta_l^0| \leq \frac{\Delta}{T^{1/2}} \text{ y } r \in [\underline{r}, \bar{r}]\}$. Para (β, r) dado, el proceso convergerá a 0. Por lo tanto, para obtener el resultado deseado, necesitamos una ULLN. Dado que B_Δ

es acotado, podemos encontrar una partición B_{ν, ℓ_j} , centrada en ϕ^{ℓ_j} , para $j = 1, \dots, J < \infty$, tal que $\sup_{\phi \in B_{\nu, \ell_j}} \varrho(\phi^{\ell_j}, \phi) < \nu$ y

$$P \left[\sup_{\phi \in B_{\Delta}} \frac{\Delta}{T} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} h_{t-j}^i(\phi^{\ell_j}) \right| > \eta \right] \leq P \left[\max_{\ell_j} \frac{\Delta}{T} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} h_{t-j}^i(\phi^{\ell_j}) \right| > \eta/2 \right] +$$

$$P \left[\max_{\ell_j} \frac{\Delta}{T} \left| \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j}(\phi^{\ell_j}, v) - E \left[|\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j}(\phi^{\ell_j}, v) \right] \right| > \eta/4 \right],$$

usando que $\frac{\Delta}{T} \sum_{t=1}^T E \left[|\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j}(\phi^{\ell_j}, v) \right] \leq \nu K < \eta/2$ para ν suficientemente pequeño, con \bar{s}_{t-j} definido al principio de la sección. Ahora,

$$P \left[\max_{\ell_j} \frac{\Delta}{T} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} h_{t-j}^i(\beta, r) \right| > \eta/2 \right] \leq \frac{\Delta^2 J^5 K}{T \eta^2} \xrightarrow{T \rightarrow \infty} 0,$$

$$P \left[\max_{\ell_j} \frac{\Delta}{T} \left| \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j} - E \left[|\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j} \right] \right| > \eta/4 \right] \leq \frac{\Delta^2 J^4 K}{T \eta^2} \xrightarrow{T \rightarrow \infty} 0,$$

dado que $|\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \bar{s}_{t-j}(\phi^{\ell_j}, v)$ es una secuencia $L_2 - mixingala$ de tamaño $-1/2$ con constante finita. Esta propiedad se obtiene utilizando las propiedades de $L_2 - NED$ de $1(\bar{A}_{i,t}(\phi^*, v))$ y $\bar{v}_t(\phi^*, v)$ establecidas en el Teorema 3.1. Por lo tanto, reagrupando resultados, para $i = 1$ tenemos que para $\forall \eta > 0$,

$$\lim_{T \rightarrow \infty} P \left[\sup_{\phi \in B_{\Delta}} \frac{\Delta}{T} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} h_{t-j}^1(\beta, r) \right| > \eta \right] = 0.$$

Ahora probaremos el mismo resultado para $i = 2$. En este caso,

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} \left(1(\hat{B}_{t-j,1}) \hat{e}_{t-j} - 1(B_{t-j,1}^0) \varepsilon_{t-j} \right) \right| \leq$$

$$\leq \sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} 1(\hat{B}_{t-j,1}) (\hat{e}_{t-j} - \varepsilon_{t-j}) \right| +$$

$$+ \sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} \left(1(\hat{B}_{t-j,1}) - 1(B_{t-j,1}^0) \right) \varepsilon_{t-j} \right|.$$

Para el primer supremo, siguiendo los mismos pasos que en el caso de $i = 1$, obtenemos que para $\forall \eta > 0$,

$$\lim_{T \rightarrow \infty} P \left[\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} 1 \left(\hat{B}_{t-j,1} \right) (\hat{e}_{t-j} - \varepsilon_{t-j}) \right| > \eta \right] = 0.$$

Para el segundo supremo, debemos probar que para $\forall \eta, \varepsilon > 0$, $\exists T_\varepsilon$ tal que para $\forall T > T_\varepsilon$

$$P \left[\sup_{\phi \in B_\Delta} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} [1 (|e_{t-j}(\beta, r)| < r) - 1 (|\varepsilon_{t-j}| < r)] \varepsilon_{t-j} \right| > \eta \right] < \varepsilon.$$

Desarrollamos la prueba para

$$P \left[\sup_{\phi \in B_\Delta} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} [1 (r < \varepsilon_{t-j} < r + v_{t-j}(\beta, \beta^0, r))] \varepsilon_{t-j} \right| > \eta \right] < \varepsilon,$$

para el otro caso, es igual. Para ello, definimos la siguiente partición del espacio paramétrico B_Δ , $B_{v,\ell_j} = \prod_{l=1}^5 B_{v,l}$, con $B_{v,l} = \{\beta_l \in B_\Delta : \beta_l \in [\beta_l^{j_l}, \beta_l^{j_l+1}]\}$, para $\beta_l^{j_l} = \beta_l^0 - \Delta T^{-1/2} + \nu T^{-1/2} j_l$ y $l = 1, \dots, 4$; y $B_{v,5} = \{r \in B_b : r \in [r^{j_5}, r^{j_5+1}]\}$ con $r^{j_5} = \underline{r} + \nu T^{-1/2} j_5$. Como en casos anteriores, $\ell_j = (j_1, j_2, j_3, j_4, j_5)$ y $\phi^{\ell_j} = (\beta_1^{j_1}, \beta_2^{j_2}, \beta_3^{j_3}, \beta_4^{j_4}, r^{j_5})$. Acotamos el supremo a través de esta partición,

$$\begin{aligned} \sup_{\phi \in B_\Delta} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} [1 (r < \varepsilon_{t-j} < r + v_{t-j}(\beta, \beta^0, r))] \varepsilon_{t-j} \right| \leq \\ \max_{\ell_j} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} [1 (r^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + v_{t-j}^{\ell_j})] \varepsilon_{t-j} \right| + \end{aligned} \quad (3.33)$$

$$+ \max_{\ell_j} \sup_{\phi \in B_{v,\ell_j}} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \Upsilon_t(\beta^{\ell_j}, r^{\ell_j}, \beta, r) \right|, \quad (3.34)$$

$$(3.35)$$

con $\Upsilon_t(\beta^{\ell_j}, r^{\ell_j}, \beta, r) = \sum_{j=1}^t \theta^{0j-1} [1 (r^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + v_{t-j}^{\ell_j}) - 1 (r < \varepsilon_{t-j} < r + v_{t-j})] \varepsilon_{t-j}$, $v_{t-j}^{\ell_j} = v_{t-j}(\beta^{\ell_j}, \beta^0, r^{\ell_j})$ y $v_{t-j} = v_{t-j}(\beta, \beta^0, r)$. El máximo de (3.34), teniendo en cuenta que $v_{t-j}(\beta^{\ell_j}, \beta^0, r^{\ell_j}) - v_{t-j}(\beta, \beta^0, r) = v_{t-j}(\phi^{\ell_j}, \phi)$, puede ser acotado por la suma de los

siguientes tres máximos,

$$\max_{\ell_j} \frac{1}{\sqrt{T}} \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \left[1 \left(r^{\ell_j} + v_{t-j}^{\ell_j} - \bar{v}_{t-j} \left(\phi^{\ell_j}, \frac{\nu}{T^{1/2}} \right) < \varepsilon_{t-j} < r^{\ell_j} + \frac{\nu}{T^{1/2}} \right) \right] |\varepsilon_{t-j}| \quad (3.36)$$

$$\max_{\ell_j} \frac{1}{\sqrt{T}} \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \left[1 \left(r^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + \frac{\nu}{T^{1/2}} \right) \right] |\varepsilon_{t-j}| \quad (3.37)$$

$$\max_{\ell_j} \frac{1}{\sqrt{T}} \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} \left[1 \left(r^{\ell_j} + v_{t-j}^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + \frac{\nu}{T^{1/2}} + v_{t-j}^{\ell_j} + \bar{v}_{t-j} \left(\phi^{\ell_j}, \frac{\nu}{T^{1/2}} \right) \right) \right] |\varepsilon_{t-j}| \quad (3.38)$$

A continuación probaremos que para $\forall \eta, \varepsilon > 0$, $\exists T_\varepsilon$ tal que para $\forall T > T_\varepsilon$,

$$P \left[\max_{\ell_j} \frac{1}{\sqrt{T}} \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}) \geq \eta \right] < \varepsilon,$$

con

$$U_t(\phi^{\ell_j}) = \left[1 \left(r^{\ell_j} + v_t^{\ell_j} < \varepsilon_t < r^{\ell_j} + \frac{\nu}{T^{1/2}} + v_t^{\ell_j} + \bar{v}_t \left(\phi^{\ell_j}, \frac{\nu}{T^{1/2}} \right) \right) \right] |\varepsilon_t|,$$

y siguiendo los mismos pasos, se obtiene el mismo resultado para los términos (3.37) y (3.38), y por lo tanto,

$$P \left[\max_{\ell_j} \sup_{\phi \in B_{v, \ell_j}} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \Upsilon_t(\beta^{\ell_j}, r^{\ell_j}, \beta, r) \right| \geq \eta \right] < \varepsilon.$$

Para obtener dicho resultado, usaremos que, como se probó a lo largo del Teorema 3.1, $v_{t-j}^{\ell_j}$ y $\bar{v}_{t-j}(\phi^{\ell_j}, \frac{\nu}{T^{1/2}})$ son $L_2 - NED$ de tamaño $-\infty$ con constante $T^{-1/4}K$. Definimos,

$$U_t^m(\phi^{\ell_j}) = \left[1 \left(r^{\ell_j} + v_t^{\ell_j, m} < \varepsilon_t < r^{\ell_j} + \frac{\nu}{T^{1/2}} + v_t^{\ell_j, m} + \bar{v}_t^m \left(\phi^{\ell_j}, \frac{\nu}{T^{1/2}} \right) \right) \right]$$

con \bar{v}_t^m y $v_t^{\ell_j, m}$ la aproximación $\mathfrak{S}_{t-m}^t - medible$ de \bar{v}_t y $v_t^{\ell_j}$ respectivamente. Por lo tanto,

$$\begin{aligned} \frac{1}{\sqrt{T}} \sum_{t=1}^T |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} U_{t-j}(\ell_j) &= \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) + \\ &+ \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} [U_{t-j}(\phi^{\ell_j}) - U_{t-j}^{m-j}(\phi^{\ell_j})] + \\ &+ \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=m}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}) + \frac{1}{\sqrt{T}} \sum_{t=1}^{m-1} |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}). \end{aligned}$$

Ahora, probaremos que para todo $\eta > 0$ y $\varepsilon > 0$, $\exists m, T < \infty$ suficientemente grandes, tal que

$$P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=1}^{m-1} |\varepsilon_t| \sum_{j=1}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}) \right| \geq \eta \right] < \varepsilon \quad (3.39)$$

$$P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} [U_{t-j}(\phi^{\ell_j}) - U_{t-j}^{m-j}(\phi^{\ell_j})] \right| \geq \eta \right] < \varepsilon \quad (3.40)$$

$$P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=m}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}) \right| \geq \eta \right] < \varepsilon \quad (3.41)$$

$$P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) \right| \geq \eta \right] < \varepsilon. \quad (3.42)$$

Para (3.39), la prueba es directa. Para la segunda probabilidad, (3.40), usaremos que $\Gamma_{t,1}^m = |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} [U_{t-j}(\phi^{\ell_j}) - U_{t-j}^{m-j}(\phi^{\ell_j})]$ es una secuencia L_2 -mixingala de tamaño $-1/2$ con constante $m^{1/2} \lambda^{\frac{m-1}{2}} T^{-1/4} K$. Este resultado se obtiene fácilmente por el Teorema 2.12, teniendo en cuenta que $\Gamma_{t,1}^m$ es $L_2 - NED$ de tamaño $-\infty$ con constante $\|\Gamma_{t,1}^m\|_2 \leq m^{1/2} \lambda^{\frac{m-1}{2}} T^{-1/4} K$. Por lo tanto, usando las propiedades de los procesos L_2 -mixingalas,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T [\Gamma_{t,1}^m - E(\Gamma_{t,1}^m)] \right| \geq \eta/2 \right] &\leq \sum_{j=1}^{\frac{2\Delta}{\nu}} \frac{j^3 T T^{1/2} m \lambda^{m-1} T^{-1/2} K}{T \eta^2} \leq \frac{m \lambda^{m-1} K}{\eta^2 \nu^4} \leq \varepsilon \\ \max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T [E(\Gamma_{t,1}^m)] \right| &\leq T^{1/2} m \lambda^{m-1} T^{-1/2} K \leq m \lambda^{m-1} K \leq \eta/2, \end{aligned}$$

tomando m suficientemente grande.

Para la tercera probabilidad, (3.41), como en el caso anterior, usaremos que $\Gamma_{t,2}^m = |\varepsilon_t| \sum_{j=m}^t \theta^{0j-1} [U_{t-m-j}(\ell_j)]$ es un proceso L_2 - *mixingala* de tamaño $-1/2$ con constante $T^{-1/4}K$, con lo que

$$\begin{aligned} \max_{\ell_j} |\theta^{0m-1}| \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T [E(\Gamma_{t,2}^m)] \right| &\leq |\theta^{0m-1}| K \leq \eta/2 \\ P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T |\varepsilon_t| \sum_{j=m}^t \theta^{0j-1} U_{t-j}(\phi^{\ell_j}) \right| \geq \eta \right] &\leq P \left[\max_{\ell_j} \left| \frac{|\theta^{0m-1}|}{\sqrt{T}} \sum_{t=m}^T [\Gamma_{t,2}^m - E(\Gamma_{t,2}^m)] \right| \geq \eta/2 \right] \\ &\leq |\theta^{0m-1}| \sum_{j=1}^{\frac{2\Delta}{\nu}} \frac{j^3 K}{\eta^2} \leq \varepsilon, \end{aligned}$$

para m suficientemente grande.

Finalmente, para la cuarta probabilidad, (3.42), definimos $\Gamma_{t,3} = |\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) - E \left[|\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) \right]$, que cumplirá que,

$$\begin{aligned} E \left[|\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) \right] &\leq \frac{\nu}{T^{1/2}} K \\ E \left(\sum_{t=m}^T \Gamma_{t,3} \right)^4 &\leq m^3 [T\nu^2 + T^{1/2}\nu] K, \end{aligned}$$

usando que $E(\Gamma_{t,3}/\mathfrak{S}_{t-m-1}^t) = 0$. Por lo tanto,

$$\begin{aligned} P \left[\max_{\ell_j} \left| \frac{1}{\sqrt{T}} \sum_{t=m}^T \Gamma_{t,3} \right| \geq \eta/2 \right] &\leq \sum_{j=1}^{\frac{2\Delta}{\nu}} \frac{j^3 T^{1/2} m^3 [T\nu^2 + T^{1/2}\nu] K}{T^2 \eta^4} \leq \frac{m^3 b^4 K}{T^{1/2} \eta^4 \nu^3} \xrightarrow{T \rightarrow \infty} 0 \\ \max_{\ell_j} \frac{1}{\sqrt{T}} \sum_{t=m}^T E \left[|\varepsilon_t| \sum_{j=1}^{m-1} \theta^{0j-1} U_{t-j}^{m-j}(\phi^{\ell_j}) \right] &\leq \nu K \leq \eta/2, \end{aligned}$$

para ν suficientemente pequeño, lo cual prueba (3.42). Probados los resultados de (3.39) a (3.42), queda probado (3.36). Siguiendo los mismos pasos, se prueban (3.37) y (3.38), por lo que para el término (3.34), se tiene que

$$P \left[\max_{\ell_j} \sup_{\phi_u} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \Upsilon_t(\beta^{\ell_j}, r^{\ell_j}, \beta, r) \right| > \eta \right] \leq \varepsilon.$$

Por último, falta obtener el mismo resultado para el término (3.33). Para ello, definimos $\Gamma_{t,4} = \sum_{j=1}^t \theta^{0j-1} \left[1 \left(r^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + v_{t-j} \left(\beta^{\ell_j}, \beta^0, r^{\ell_j} \right) \right) \right] \varepsilon_{t-j}$. Dado que $\varepsilon_t \Gamma_{t,4}$ es una *mds*, aplicando la desigualdad de Burkholder, se tiene que

$$E \left(\sum_{t=n}^T \varepsilon_t \Gamma_{t,4} \right)^4 \leq K E \left(\sum_{t=n}^T \varepsilon_t^2 \Gamma_{t,4}^2 \right)^2.$$

Como en los casos anteriores, $\Gamma_{t,4}$ es L_2 -*NED* de tamaño $-\infty$ y constante proporcional a $T^{-1/4}v^{1/2}$. Usando los Teoremas 2.12 y 2.14, se tiene que $\Gamma_{t,4}^2$ es una secuencia L_2 -*mixingala* de tamaño $-1/2$ con constante proporcional a $T^{-1/4}v^{1/2}$. Esto garantiza que,

$$P \left[\max_{\ell_j} \frac{1}{\sqrt{T}} \left| \sum_{t=1}^T \varepsilon_t \sum_{j=1}^t \theta^{0j-1} \left[1 \left(r^{\ell_j} < \varepsilon_{t-j} < r^{\ell_j} + v_{t-j} \left(\beta^{\ell_j}, \beta^0, r^{\ell_j} \right) \right) \right] \varepsilon_{t-j} \right| > \eta \right] \leq \varepsilon.$$

La prueba es similar a la seguida para el término (2.62) en la prueba del lema 2.14 en el Anexo 2 del Capítulo 2. Esto finaliza la prueba del lema. ■

Lema 3.11 *Bajo los supuestos del Teorema 3.6,*

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{T} \sum_{t=1}^T H_{t-1}^{i,l} \left(\widehat{\beta}(r) \right) - H_{t-1}^{i,l} \left(\beta^0, r \right) = o_p(1).$$

Prueba del Lema 3.11. Usando los mismos argumentos que en el lema 3.10, la prueba del lema se sigue de

$$P \left[\sup_{\phi \in B_\Delta} \frac{1}{T} \sum_{t=1}^T \left(H_{t-1}^{i,l} \left(\beta, r \right) - H_{t-1}^{i,l} \left(\beta^0, r \right) \right) > \eta \right] < \varepsilon,$$

para todo $\eta > 0$, $\varepsilon > 0$ y T suficientemente grande. De la definición de $H_{t-1}^{i,l}(\beta, r)$ y $H_{t-1}^{i,l}(\beta^0, r)$,

$$\begin{aligned} \left(H_{t-1}^{i,l}(\beta, r) - H_{t-1}^{i,l}(\beta^0, r) \right) &= \left(h_t^{(i)}(\beta, r) - h_t^{(i)}(\beta^0, r) \right) h_t^{(l)}(\beta, r) + \\ &\quad + \left(h_t^{(l)}(\beta, r) - h_t^{(l)}(\beta^0, r) \right) h_t^{(i)}(\beta^0, r). \end{aligned}$$

Dado que para $i = 3, 4$, $h_t^{(i)}(\beta, r)$ no depende de (β, r) , y dado que $e_t(\beta, r) - \varepsilon_t = \sum_{i=1}^4 (\beta_i - \beta_i^0) h_t^{(i)}(\beta, r)$, el término de interés será

$$\sup_{\phi \in B_\Delta} \frac{1}{T} \sum_{t=1}^T h_{t-1}^{(i)}(\beta^0, r) \sum_{j=1}^t \theta^{0j-1} [1(|e_{t-j}(\beta, r)| < r) - 1(|\varepsilon_{t-j}| < r)] \varepsilon_{t-j},$$

para $i = 1, \dots, 4$. La prueba se obtendrá de la misma forma que en el lema 3.10 para $i = 2$, teniendo en cuenta que en este caso el proceso está multiplicado por $h_{t-1}^{(i)}(\beta^0, r)$, que es $L_2 - NED$ de tamaño ∞ y constante finita y la suma está estandarizada por T y no $T^{1/2}$.

■

Por lo tanto, usando los lemas 3.10 y 3.11, tenemos que

$$\begin{aligned} Q_T(\hat{\beta}, r) - Q_T(\beta^0, r) &= \sum_{i=1}^4 \sum_{l=1}^4 \sum_{t=1}^T (\hat{\beta}_i - \beta_i^0) H_{t-1}^{i,l}(\beta^0, r) (\hat{\beta}_l - \beta_l^0) + \\ &\quad + 2 \sum_{t=1}^T \varepsilon_t \sum_{i=1}^4 (\hat{\beta}_i - \beta_0) h_t^{(i)}(\beta^0, r) + o_p(1) \end{aligned}$$

con el $o_p(1)$ uniforme en r . Finalmente, de forma paralela al caso observable, probaremos que $T^{-1} \sum_{t=1}^T H_{t-1}^{i,l}(\theta^0, r)$ converge uniformemente en r a $H^{i,l}(r) = E(h_t^{(i)}(\theta^0, r) h_t^{(l)}(\theta^0, r))$, y que $s_T(r) = T^{-1/2} \sum_{t=1}^T \varepsilon_t h_t(\beta^0, r)$ es tight en r para $i = 1, \dots, 4$. Esto se establece en los siguientes dos lemas.

Lema 3.12 *Bajo los supuestos del Teorema 3.6*

$$\sup_{r \in [\underline{r}, \bar{r}]} \frac{1}{T} \sum_{t=1}^T [H_{t-1}^{i,l}(\beta^0, r) - H^{i,l}(r)] = o_p(1),$$

con $E(H_{t-1}^{i,l}(\beta^0, r)) = H^{i,l}(r)$.

Prueba del Lema 3.12. La prueba se deriva directamente del lema 2.13 del Capítulo 2, dado que en este caso los procesos son idénticos con $z_t = \varepsilon_t$, que cumple las condiciones del lema. ■

Lema 3.13 *Bajo los supuestos del Teorema 3.6, el proceso $s_T(r) = T^{-1/2} \sum_{t=1}^T \varepsilon_t h_t(\beta^0, r)$ es tight.*

Prueba del Lema 3.13. Como en el lema anterior, el proceso es igual al del lema 2.14 en el Capítulo 2, con $z_t = \varepsilon_t$, que cumple las condiciones del lema, y por tanto este lema sale de la aplicación directa del lema 2.14. ■

Prueba del Teorema 3.6. Dados los resultados anteriores, se tiene que $T^{1/2} \left(\hat{\beta}(r) - \beta^0 \right) = H_T^{-1}(\beta^0, r) s_T(r) + o_p(1)$ con el $o_p(1)$ uniforme en r . De nuevo, los procesos $H_T^{-1}(\beta^0, r)$ y $s_T(r)$ cumplen las mismas propiedades que los definidos para el Teorema 2.6 en el Capítulo 2, por tanto, la prueba del Teorema 3.6 sigue los mismos pasos que la prueba del Teorema 2.6. ■

Capítulo 4

Método No Lineal de Descomposición en Transitorio y Permanente.

Resumen: *Este capítulo propone una nueva descomposición permanente-transitorio no lineal, que comparte la ortogonalidad de componentes inobservados y el espíritu de Beveridge-Nelson. Esta descomposición utiliza un nuevo esquema de identificación, que se deriva de los modelos ARTIMA, y basado en dos aspectos: (i) el efecto de la perturbación en el largo plazo (permanente o transitorio) y (ii) el tamaño de la perturbación (grande o pequeña). Como vimos, a partir de los modelos ARTIMA es posible desarrollar una estrategia de contraste para determinar si las perturbaciones grandes y pequeñas tienen un efecto diferente en el largo plazo, así como si una de ellas es puramente transitoria, lo cual permite contrastar la idoneidad de la descomposición propuesta. También se obtiene la descomposición de Beveridge-Nelson no lineal para los modelos ARTIMA, probándose que su comportamiento se encuentra entre el de la descomposición lineal y la ortogonal de componentes inobservados. El Capítulo presenta un nuevo contraste, de fácil aplicación, para la existencia de asimetrías en la persistencia en función del tamaño de la perturbación. Los resultados del capítulo se aplicarán a un índice de precios, para medir la calidad del mercado, y al PNB, para analizar si la persistencia de las perturbaciones es asimétrica.*

4.1. Introducción

¿Cuales son las perturbaciones que guían la economía? ¿De donde procede la persistencia presente en la mayoría de las series temporales económicas? ¿Que proporción de la variación de la producción es debida a perturbaciones de la demanda o la oferta? Todos los intentos de responder estas preguntas pasan por la tarea crucial de identificar las perturbaciones.

Tradicionalmente, esto se ha hecho mediante la descomposición en los componentes permanentes y transitorios inobservados de la variable analizada (ver Watson, 1986). Algunos ejemplos son: (1) La descomposición de Beveridge y Nelson (1981), ($B - N$ en adelante), donde el componente permanente es un paseo aleatorio y existe una correlación perfecta entre las perturbaciones transitorias y permanentes, debido a que solo hay una perturbación; (2) Los modelos de Componentes Inobservados ($UC - 0$) (ver Harvey, 1985, y Clark, 1987), donde el componente permanente se asume que es un paseo aleatorio pero las innovaciones están incorreladas con las del componente transitorio. Aquí, es conveniente hacer notar que no todos los modelos ARIMA pueden ser representados a través de un modelo $UC - 0$. Este problema se resuelve si eliminamos la restricción de que ambos componentes estén incorrelados, o que el componente permanente sea un paseo aleatorio. En ambas situaciones tendríamos un problema de identificación extra, que en el caso de eliminación de la restricción de paseo aleatorio en el componente permanente, es resuelto imponiendo una condición ad-hoc de suavidad en dicho componente (ver la descomposición “canónica” en Pierce, 1979). Todas estas descomposiciones presentan dos problemas adicionales a estos recogidos en Quah (1992), y que son: (i) en todo periodo t hay siempre una perturbación permanente, y (ii) ningún supuesto de los necesarios en la descomposición permanente y transitorio son contrastables. La aproximación propuesta en este capítulo trata de resolver la cuestión de la identificación evitando estos problemas.

Nuestra proposición esta basada en dos pilares. El primero, como fue mencionado ante-

riormente, en la aceptación de la existencia de perturbaciones permanentes y transitorias, y el segundo, que estas se comportan de forma diferente en el largo plazo dependiendo de sus propias características (signo, tamaño, estado de la economía, etc.). La primera premisa, como es bien sabido (ver Quah, 1992), no es suficiente para identificar las perturbaciones permanentes y transitorias de una serie univariante. Este capítulo muestra que con nuestro segundo supuesto, no solo la identificación de las perturbaciones permanentes y transitorias de una variable económica simple es factible, si no que también se puede contrastar la validez del supuesto utilizado. Con la intención de implementar nuestra proposición usaremos los modelos threshold introducidos en los capítulos anteriores, los modelos ARTIMA. Como vimos, la variable threshold puede representar cualquier característica de la economía o las perturbaciones (grandes o pequeñas, positivas o negativas, expansión o recesión, etc.) en la que estemos interesados. Este diseño threshold es capaz de capturar la posibilidad de que diferentes características de las perturbaciones generen diferentes efectos a largo plazo. Como hemos visto, bajo determinadas condiciones, la característica asociada al régimen threshold donde la media móvil presenta una raíz unitaria implica que la perturbación tendrá un efecto puramente transitorio.

En este capítulo nos centraremos en el tamaño de la perturbación, como en el modelo STOPBREAK de Engle y Smith (1999). Sin embargo, el análisis que se lleva a cabo en este capítulo es extensible a otros casos o características tratadas en los capítulos anteriores. La hipótesis que mantendremos a lo largo del Capítulo será que las perturbaciones grandes provocarán efectos permanentes mientras que las pequeñas producirán solo efectos transitorios. Esta hipótesis se basa en el supuesto de que la probabilidad de que ciertas series temporales reviertan a su media, después de una perturbación grande, es inferior al caso de una perturbación pequeña. Por ejemplo, en el mercado de activos, los inversores pueden percibir que las perturbaciones grandes contienen información mientras que las perturbaciones pequeñas son mero ruido. En macroeconomía, una cuestión no resuelta es la de si

las fluctuaciones económicas se deben a la acumulación de pequeñas perturbaciones transitorias o por el contrario, provienen de perturbaciones de gran tamaño infrecuentes. Un aspecto importante de nuestra proposición es la posibilidad de contrastar estas hipótesis utilizadas.

La identificación de las perturbaciones permanentes y transitorias a través de los modelos ARTIMA, nos permitirá construir una nueva descomposición permanente-transitorio no lineal de la serie original. En esta nueva descomposición, el componente permanente estará formado sólo por las perturbaciones permanentes, y el transitorio, sólo por las transitorias. Dadas las características de los modelos ARTIMA, esto garantiza que ambos componentes sean ortogonales. También es posible obtener una descomposición $B - N$ no lineal basada en el modelo ARTIMA, que como se probará a lo largo del capítulo, muestra un comportamiento a medio camino entre la descomposición de $B - N$ lineal y la de $UC - 0$.

El capítulo finaliza con dos aplicaciones del modelo ARTIMA-Shock-Size y su identificación de las perturbaciones: primero, a los precios de los activos, donde siguiendo a Hasbrouck (1993), consideramos el tamaño del componente transitorio como una medida de la calidad o eficiencia de los mercados de activos; y segundo al PNB, donde la mayoría de las investigaciones respecto del tamaño de su componente permanente ha impuesto implícitamente la simetría, y en los escasos casos en los que se permiten las asimetrías, no se ha alcanzado una conclusión definitiva (ver Elwood, 1998).

El resto del capítulo se estructura de la siguiente manera. En la siguiente sección presentamos un nuevo estadístico para el contraste de la hipótesis de efecto threshold. En la Sección 3 definimos dos nuevas descomposiciones permanente-transitorio que se comparan con las descomposiciones lineales ya existentes. Las Secciones 4 y 5 presentan dos aplicaciones del modelo ARTIMA-Shock-Size, una, con el objetivo de medir la eficiencia de los mercados de activos, y la otra, con el de analizar si la persistencia de las perturbaciones en el PNB es asimétrica. Finalmente, la Sección 7 presenta las conclusiones. Las pruebas se

encuentran en el Anexo al final del capítulo.

4.2. Contraste de Existencia de Asimetrías

Como hemos visto, un aspecto importante del método que se propone en este Capítulo es la posibilidad de contrastar el supuesto de identificación basado en el comportamiento asimétrico de las perturbaciones. Tanto en el Capítulo 2 como en el 3, se presentan los estadísticos necesarios para este contraste, en el Capítulo 2, para el caso observable, y en el Capítulo 3, para el caso de la propia perturbación. En concreto, la estrategia propuesta en estos capítulos consistirá en contrastar, en una primera etapa, la existencia de asimetrías, y en una segunda, cuando se acepte la existencia de asimetrías, la existencia de perturbaciones puramente transitorias. Los estadísticos para el primer contraste, propuestos en ambos capítulos, se construyen desde la hipótesis alternativa, esto es, desde el modelo no lineal. En esta sección presentamos un nuevo estadístico construido desde la hipótesis nula, y por lo tanto acorde con los contrastes del multiplicador de Lagrange. En este sentido, el contraste es similar al propuesto por Engle y Smith (1999) para el modelo STOPBREAK o el propuesto por Brännäs, De Gooijer y Teräsvirta (1998). Este tipo de estadísticos presentan la ventaja de que su cálculo es más fácil, ya que el estadístico se construye bajo el modelo lineal, aunque su potencia suele ser inferior.

Para obtener el estadístico bajo la hipótesis nula de linealidad, ajustaremos a y_t , la serie de interés, un modelo ARIMA(p,d,q),

$$\Phi_p(L) (1 - L)^d y_t = \mu + \Theta_q(L) \varepsilon_t, \quad (4.1)$$

y usaremos los residuos, $\hat{\varepsilon}_t$, para construir la siguiente regresión auxiliar

$$\hat{\varepsilon}_t = \alpha(r) \hat{\varepsilon}_{t-1} 1(|\hat{\varepsilon}_{t-1}| > r) + u_t. \quad (4.2)$$

Bajo la hipótesis nula α será igual a 0, y bajo la alternativa, $\alpha \neq 0$. Esta hipótesis puede ser contrastada a través de un t-ratio estándar en la ecuación (4.2). El único problema es

que bajo la nula, r no estará identificado, por lo tanto, como se propuso en los Capítulos 2 y 3, procedemos como Davis (1977, 1987), definiendo el estadístico como el supremo del t -ratio sobre todos los valores de $r \in [\underline{r}, \bar{r}]$. La distribución asintótica se presenta en el siguiente Teorema, para el que es necesario el siguiente supuesto:

A.0 ε_t iid $(0, \sigma_\varepsilon)$, con función de densidad $0 < f(\varepsilon_t) < \infty$ uniformemente continua para $\forall \varepsilon_t$, y $\|\varepsilon_t\|_{2\gamma} < \infty$ con $\gamma > 2$,

y donde \Rightarrow implica convergencia uniforme en distribución. Antes de enunciar el Teorema, denotemos por $\hat{\lambda} = (\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)$ al estimador de Mínimos Cuadrados de $\lambda^0 = (\phi_1^0, \dots, \phi_p^0, \theta_1^0, \dots, \theta_q^0)$ en la ecuación (4.1), entonces, bajo A.0 es fácil probar que,

$$T^{1/2}(\hat{\lambda}_i - \lambda_i^0) = T^{1/2} \sum_{t=1}^T \varepsilon_t s_{i,t-1}(\lambda^0) H_{i,i}^{-1}(\lambda^0) + o_p(1), \quad \text{para } i = 1, \dots, p+q,$$

tal que,

$$T^{1/2}(\hat{\lambda} - \lambda^0) \xrightarrow{d} N(\vec{0}, M_\lambda).$$

Teorema 4.1 Sea $t_{\alpha(r)}$ el t -ratio para el parámetro $\alpha(r)$ en la regresión auxiliar (4.2) para un r dado. Bajo el supuesto A.0,

$$t_{\alpha(r)} \Rightarrow D(r) \quad \text{en } [\underline{r}, \bar{r}],$$

con

$$D(r) \equiv \left[D_1(r) + \sum_{i=1}^{p+q} V_{\lambda_i} \Sigma_\varepsilon(r) \right] [\Sigma_\varepsilon(r) \sigma_\varepsilon^2]^{-1/2},$$

donde $D_1(r)$ es un proceso gaussiano continuo con función de covarianzas igual a

$$M_1(r_1, r_2) = E(\varepsilon_t^2) E(\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| < \min(r_1, r_2))),$$

V_{λ_i} es una v.a normal con distribución dada por la distribución asintótica de $T^{1/2}(\hat{\lambda}_i - \lambda_i^0)$ y

$\Sigma_\varepsilon(r) = E(\varepsilon_t^2 1(|\varepsilon_t| < r))$. Finalmente, $Cov(D_1(r), V_{\lambda_i}) = E(\varepsilon_t^2) E(\varepsilon_{t-1} s_{i,t-1}(\lambda^0)) H_{i,i}^{-1}(\lambda^0)$.

Por lo tanto, usando el Teorema del funcional continuo,

$$\sup_{r \in [\underline{r}, \bar{r}]} |t_{\alpha(r)}| \xrightarrow{d} \sup_{r \in [\underline{r}, \bar{r}]} D(r).$$

Esta distribución asintótica $\sup_r D(r)$ dependerá de parámetros ruidosos, con lo que no será posible su tabulación. Así, para obtener el p – valor del contraste se propone un método Wild Bootstrap, que describimos a continuación.

Denotemos $\widehat{\varepsilon}_t, \widehat{\Phi}_p(L), \widehat{\mu}$ y $\widehat{\Theta}_q(L)$ a los estimadores de MCO de $\varepsilon_t, \Phi_p, \mu$ y $\Theta_q(L)$ respectivamente de la ecuación (4.1). Entonces, para $i = 1, \dots, M$ ejecutamos los siguientes pasos:

- (i) generamos $\{\eta_t^i\}_{t=1}^T$ como v.a iid $N(0, 1)$;
- (ii) obtenemos $v_t^i = \eta_t^i \widehat{\varepsilon}_t$ e y_t^i de $\widehat{\Phi}_p(L) (1 - L) y_t^i = \widehat{\mu} + \widehat{\Theta}_q(L) v_t^i$;
- (iii) estimamos por MCO $\{\widetilde{\varepsilon}_t^i\}_{t=1}^T$ en la ecuación (4.1) para $\{y_t^i\}_{t=1}^T$;
- (iv) estimamos por MCO $\widetilde{\alpha}^i(r)$ en la ecuación (4.2) para $\{\widetilde{\varepsilon}_t^i\}_{t=1}^T$;
- (v) calculamos $\sup_r^i |t_{\widetilde{\alpha}(r)}|$.

Por lo tanto, tomando la distribución empírica del $\sup_r^i |t_{\widetilde{\alpha}(r)}|$ como una aproximación de la distribución asintótica del $\sup_r |t_{\alpha(r)}|$, rechazaremos la hipótesis nula al nivel ϵ si

$$\sup_r |t_{\alpha(r)}| > q_\epsilon,$$

donde $\epsilon = \widehat{P}_{\widetilde{\alpha}} [\sup_r^i |t_{\widetilde{\alpha}(r)}| > q_\epsilon]$, y $\widehat{P}_{\widetilde{\alpha}}$ la distribución empírica de $\sup_r^i |t_{\widetilde{\alpha}(r)}|$.

Si la hipótesis nula de linealidad es rechazada, entonces procederemos a contrastar la existencia de perturbaciones transitorias mediante la construcción del t-ratio de $\alpha_4 = 1$ en la ecuación (3.2) y según se propone en la Sección 3.4.2 del Capítulo 3. De acuerdo al Teorema 3.5, este estadístico seguirá una distribución normal estándar.

Finalmente, el resultado del Teorema 4.1 es fácilmente extensible al resto de los casos considerados en la Tesis. Para ello, $\sup_r |t_{\tilde{\alpha}(r)}|$ se obtendrá de la siguiente ecuación de regresión auxiliar general:

$$\hat{\varepsilon}_t = \alpha(r)\hat{\varepsilon}_{t-1}1(z_{t-1} > r) + u_t.$$

4.2.1. Comportamiento del contraste en muestras finitas

En esta sección analizamos el comportamiento, en términos de tamaño y potencia, del contraste propuesto en la sección anterior. En el estudio del tamaño consideraremos los modelos ARIMA(0,1,1) generados en los ejercicios de simulación de los Capítulos 3 y 4, esto es:

$$(1 - L)y_t = \varepsilon_t - \theta\varepsilon_{t-1}$$

con $\theta = (-0.5, 0, 0.5)$. Para la potencia, generaremos los procesos ARTIMA-Shock-Size del Capítulo 3,

$$(1 - L)y_t = \varepsilon_t - \theta_1\varepsilon_{t-1}1(|\varepsilon_{t-1}| > r) - \theta_2\varepsilon_{t-1}1(|\varepsilon_{t-1}| < r)$$

donde $(\theta_1, \theta_2) = (0, 1)$ y $r = (0.35, 0.5, 0.7, 0.9, 1.1, 1.3, 1.6, 2.0)$. En ambos procesos, ARIMA y ARTIMA, como en los estudios de simulación anteriores, ε_t es *iid* $N(0, 1)$, el tamaño seleccionado para el Monte Carlo es $M = 400$ y para el Wild Bootstrap $B = 400$, el tamaño muestral $T = 200$ y los niveles de significación $\alpha = 0.05, 0.10$.

En la Tabla 4.1 se recogen los resultados para el estudio del tamaño, mientras que en la Tabla 4.2 se recogen los resultados para la potencia.

Al igual que sucedía en el contraste de Wald propuesto en el Capítulo anterior, a medida que aumenta r la potencia del contraste aumenta. Sin embargo, a partir de un valor de r , (1.3 en este caso) la potencia del contraste disminuye. La razón es obvia, a medida que r se acerca a ∞ más cerca volvemos a estar de la hipótesis nula. Finalmente, comparado con

Cuadro 4.1: Tamaño del contraste lineal de existencia de threshold

θ	$\alpha - level$	
	0.05	0.10
-0.5	0.055	0.090
0.0	0.055	0.125
0.5	0.065	0.125

Cuadro 4.2: Potencia del contraste lineal de existencia de threshold

$\alpha - level$	r							
	0.35	0.5	0.7	0.9	1.1	1.3	1.6	2.0
0.05	0.145	0.275	0.485	0.565	0.570	0.530	0.515	0.435
0.10	0.205	0.370	0.625	0.705	0.720	0.765	0.690	0.625
$E(\varepsilon_t)$	0.269	0.377	0.511	0.626	0.724	0.802	0.886	0.950
$Var(\varepsilon_t^L)$	0.010	0.030	0.078	0.151	0.248	0.361	0.534	0.735
$Var(\varepsilon_t^s)$	1.032	1.013	0.965	0.892	0.794	0.682	0.509	0.308

$$\varepsilon_t^L = \varepsilon_t 1(|\varepsilon_t| > r) \text{ y } \varepsilon_t^s = \varepsilon_t - \varepsilon_t^L.$$

dicho contraste de Wald, el comportamiento en el tamaño es muy similar, aunque respecto a la potencia, y como cabía esperar, su comportamiento es ligeramente inferior.

4.3. Descomposición Permanente-Transitorio

Tanto por razones empíricas como teóricas, frecuentemente es deseable descomponer una serie temporal económica en la suma de los componentes transitorios y permanentes inobservados que generan la serie. La mayor parte de la literatura se ha centrado en la

descomposición de modelos lineales, ver Watson (1986) y más recientemente Mortley *et al.* (2001). Como estos autores ponen de manifiesto, esto implícitamente impone la restricción de que el componente transitorio (relacionado con el business cycle) es simétrico, mientras que recientes investigaciones sugieren lo contrario (ver por ejemplo, Neftci (1984), Hamilton (1989) y Beadry y Koop (1993)). En esta sección se proponen dos nuevas descomposiciones permanente-transitorio (P-T en adelante) no lineales basadas en los modelos ARTIMA-Shock-Size, aunque extensible a otros casos tratados en la Tesis y en los que un régimen de media móvil no invertible es factible.

Primero presentamos una versión modificada de la definición de descomposición P-T de Quah (1992).

Definition 4.2 *Sea Y una secuencia integrada. Una descomposición P-T para Y será un par de secuencias (Y^P, Y^T) tal que:*

- i.) Y^P es integrado e Y^T es estacionario en covarianza;*
- ii.) Y^P no posee ninguna perturbación transitoria;*
- iii.) $Var(\Delta Y^P)$ y $Var(\Delta Y^T)$ son estrictamente positivas donde $\Delta = (1 - L)$;*
- iv.) $Y_t = Y_t^P + Y_t^T$.*

Además, si

- v.) ΔY^P está incorrelado con Y^T para todo retardo.*

entonces diremos que la descomposición $P - T$ es ortogonal.

La nueva condición *ii)* garantiza que el componente permanente esta formado sólo por perturbaciones permanentes. Merece la pena comentar que la descomposición de Beveridge y Nelson no satisface la condición *v)*, dado que sus perturbaciones permanentes y transitorias están perfectamente correladas.

Dado el siguiente modelo ARTIMA-Shock-Size general, ,

$$\Phi(L)(1 - L)y_t = \mu + (1 - \theta_1 L)\varepsilon_t^L + (1 - \theta_2 L)\varepsilon_t^s, \quad (4.3)$$

con $\theta_2 = 1$, la primera nueva descomposición $P - T$ no lineal propuesta es

$$y_t^P = \mu\Phi(1)^{-1} + y_{t-1}^P + \Phi(L)^{-1}(1 - \theta_1 L)\varepsilon_t^L, \quad (4.4)$$

$$y_t^T = \Phi(L)^{-1}\varepsilon_t^s, \quad (4.5)$$

donde el componente permanente está formado sólo por las perturbaciones permanentes ($\varepsilon_t^L = \varepsilon_t 1(|\varepsilon_t| > r)$), y el componente transitorio, sólo por perturbaciones transitorias ($\varepsilon_t^s = \varepsilon_t 1(|\varepsilon_t| \leq r)$). Es obvio que esta descomposición no lineal es una descomposición ortogonal de acuerdo a la definición 4.2. Una característica importante de esta nueva descomposición P-T es que su existencia puede ser contrastada, contrastando que $\theta_2 = 1$. Además, está permite una estructura dinámica más rica en el componente permanente que la producida por el paseo aleatorio puro.

Hay situaciones donde la teoría económica sugiere que el componente permanente sea un paseo aleatorio puro (ver la siguiente aplicación a los precios de los activos). Para estos casos proponemos aplicar la descomposición de B-N directamente al modelo ARTIMA-Shock-Size. El método desarrollado por Beveridge y Nelson (1981) consiste en definir el componente permanente como la predicción de largo plazo de la serie original, o la parte que se mantiene después de que toda la dinámica transitoria ha sido eliminada. Concretamente,

$$y_t = y_t^P + y_t^T,$$

$$y_t^P = \lim_{k \rightarrow \infty} [E(y_{t+k} | y_t, y_{t-1}, \dots) - kE(y_t - y_{t-1})],$$

$$y_t^T = y_t - y_t^P.$$

Para aplicar esta definición al modelo ARTIMA-Shock-Size (4.3), primero lo reescribimos como

$$(1 - L)y_t = \mu\Phi(1)^{-1} + C_L(L)\varepsilon_t^L + C_s(L)\varepsilon_t^s,$$

donde $C_L(L) = \Phi^{-1}(L)(1 - \theta_1 L)$ y $C_s(L) = \Phi^{-1}(L)(1 - \theta_2 L)$.

Segundo, usamos los resultados algebraicos de Phillips y Solo (1992) por los cuales el polinomio de retardos $C(L)$ puede ser escrito como $C(L) = C(1) + (1 - L)\tilde{C}(L)$. Entonces, la correspondiente descomposición B-N puede ser fácilmente obtenida como

$$y_t^P = \mu\Phi(1)^{-1} + y_{t-1}^P + C_L(1)\varepsilon_t^L + C_s(1)\varepsilon_t^s, \quad (4.6)$$

$$y_t^T = \tilde{C}_L(L)\varepsilon_t^L + \tilde{C}_s(L)\varepsilon_t^s. \quad (4.7)$$

Esta descomposición constituye nuestra segunda propuesta de descomposición no lineal.

Para el caso simple de ARTIMA(0,1,1,1),

$$y_t^P = \mu + y_{t-1}^P + (1 - \theta_1)\varepsilon_t^L + (1 - \theta_2)\varepsilon_t^s,$$

$$y_t^T = -\theta_1\varepsilon_t^L - \theta_2\varepsilon_t^s.$$

Cuando las perturbaciones pequeñas sean transitorias ($\theta_2 = 1$) la descomposición $B - N$ no lineal es

$$y_t^P = \mu + y_{t-1}^P + (1 - \theta_1)\varepsilon_t^L, \quad (4.8)$$

$$y_t^T = -\theta_1\varepsilon_t^L - \varepsilon_t^s. \quad (4.9)$$

En este caso, y cuando $\theta_1 = 0$, ambas descomposiciones no lineales coincidirán. Además, comparten la propiedad de que no en todo periodo t existe una perturbación permanente. En el caso de la descomposición $B - N$ no lineal, las perturbaciones de los componentes permanente y transitorio no están perfectamente correladas, de manera que esta descomposición de $B - N$ no lineal elimina las dos principales desventajas de la descomposición B-N lineal (otros ejemplos de descomposición $B - N$ no lineal pueden encontrarse en Clarida y Taylor, 2003). Así, la principal ventaja de esta descomposición B-N no lineal, como el siguiente Teorema prueba, no se deriva de su componente permanente sino del transitorio.

El siguiente Teorema establece la relación entre las descomposiciones de B-N y $UC - 0$

aplicadas a la representación de Wold (1938) de un proceso ARTIMA-Shock-Size $(0,1,1,1)$ y la descomposición no lineal de $B - N$ ((4.8)-(4.9)).

Teorema 4.3 *Sea y_t un proceso ARTIMA-Shock-Size $(0,1,1,1)$, entonces*

$$V^{BN}(\Delta y_t^P) = V^{ARTIMA}(\Delta y_t^P) = V^{UC-0}(\Delta y_t^P),$$

$$V^{BN}(y_t^T) \leq V^{ARTIMA}(y_t^T) \leq V^{UC-0}(y_t^T),$$

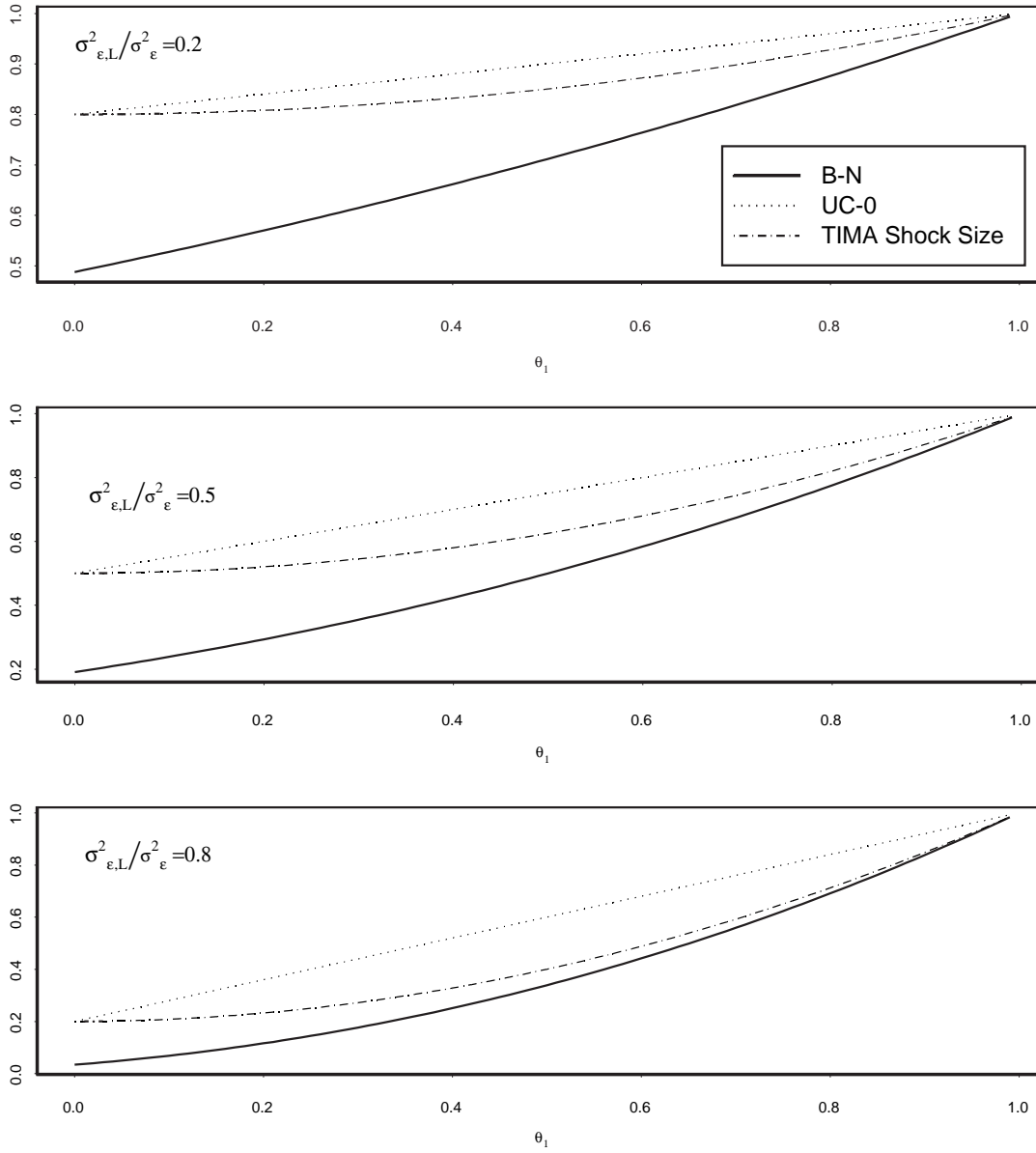
donde V^{BN} y V^{UC-0} son las varianzas de los componentes de la descomposición de $B - N$ y $UC - 0$ aplicadas a la representación de Wold de $(1 - L)y_t$, y V^{ARTIMA} es la varianza de los componentes que forman la descomposición no lineal de $B - N$ de un proceso ARTIMA-Shock-Size $(0,1,1,1)$. La desigualdad estricta se alcanzará si $0 < |\theta_1| < 1$.

Como es bien sabido, la varianza de las innovaciones del componente paseo aleatorio son siempre las mismas (iguales a la altura de la densidad espectral de $(1 - L)y_t$ en la frecuencia cero), independientemente de la descomposición utilizada. Sin embargo, la varianza del componente transitorio de cada descomposición es diferente. En particular, la varianza del componente transitorio de la descomposición $B - N$ no lineal para los ARTIMA-Shock-Size $(0,1,1,1)$ estará entre las correspondientes varianzas de la descomposición $UC - 0$ y la $B - N$ lineal de su representación de Wold. Este hecho tiene importantes consecuencias para las aplicaciones analizadas en las siguientes dos secciones.

En la Figura (4.1) se ilustra gráficamente las desigualdades del Teorema 4.3 para $\theta_2 = 1$ y diferentes valores de $\theta_1 \in (0, 1)$ y $\frac{\sigma_{\varepsilon, L}^2}{\sigma_{\varepsilon}^2}$.

Figura 4.1:

Varianza del Componente Transitorio de las diferentes descomposiciones P-T de un
proceso ARTIMA-Shock-Size



Nota: $\sigma^2_{\epsilon,L} = E(\epsilon^2 1(|\epsilon| > r))$ y $\sigma^2_{\epsilon} = E(\epsilon^2)$. $\theta_2 = 1$ y $\epsilon_t \equiv N(0, 1)$.

4.4. Aplicación a los Precios de Activos: Una medida de la Calidad del Mercado de Activos

En esta sección aplicamos los modelos ARTIMA-Shock-Size para medir las desviaciones entre los precios de transacción y el precio eficiente implícito, siguiendo la metodología desarrollada por Hasbrouck (1993). La medida de esta diferencia aparece en el análisis de los mercados financieros en dos importantes contextos. Primero, con el objetivo de evaluar la actuación de los brokers, y segundo, en el análisis comparativo de las estructuras reguladoras de los mercados. La aproximación de Hasbrouck se basa en la descomposición del precio de transacción, p_t , en dos componentes, el precio eficiente, m_t , y en el precio de error, s_t , y en la medida del tamaño relativo de s_t . Concretamente, Hasbrouck considera el logaritmo del precio de transacción observado en el periodo t , como la suma de dos componentes:

$$p_t = m_t + s_t.$$

El elemento m_t (precio eficiente) se asume que sigue un paseo aleatorio,

$$m_t = m_{t-1} + \varepsilon_t,$$

y s_t (precio de error) es un proceso estacionario en covarianza y de media cero que se modeliza como

$$s_t = \beta \varepsilon_t + \eta_t,$$

donde η_t está incorrelado con ε_t . Dado que el componente permanente esta identificado con el paseo aleatorio, la descomposición P-T correcta dependerá sólo de los supuestos sobre la estructura de autocorrelación entre ambos componentes. Si las perturbaciones de ambos componentes están perfectamente correladas ($\eta_t = 0$), usaremos la descomposición de B-N. Si ambas perturbaciones están incorreladas, ($\beta = 0$), se usará la descomposición $UC - 0$. Desafortunadamente, ninguna de las dos restricciones de identificación es contrastable,

y como se mostró en la anterior sección, ambas implican una dispersión del precio de error diferente. Además, la teoría económica no argumenta de forma definitiva en favor de ninguna de las dos restricciones de identificación.

El modelo de bid-ask spread de Roll (1984) corresponde al caso $\beta = 0$ y $\eta_t = \pm(\text{spread})/2$. Roll prueba que el bid-ask spread genera una autocorrelación de primer orden diferente de 0 en el incremento de los precios de transacción, y esta autocorrelación explica la existencia del precio de error. La restricción de correlación perfecta es un caso particular del modelo de Glosten (1987). En este modelo, el spread se debe parcialmente a la información asimétrica revelada en la transacción. Sin información pública, se obtiene $\beta > 0$ y $\eta_t = 0$.

Partiendo de estos dos modelos, desarrollamos un modelo más completo, el cual tendrá en cuenta varias cuestiones no consideradas en ninguno de los dos casos anteriores. Asumiremos la descripción del mercado hecha por Glosten (1987), esto es, hay dos clases de agentes, los informados y los no informados, con un market maker. El primer aspecto a ser considerado es que cuando hay un coste de transacción, no toda la nueva información será trasladada al precio de transacción. Solo la nueva información que implique un beneficio superior al coste de transacción se trasladará al precio de transacción. En otras palabras, las innovaciones que dirigen el precio eficiente deben ser perturbaciones “grandes” en el precio de transacción. El segundo aspecto es que las transacciones de los agentes no informados no pueden generar grandes cambios ineficientes en el precio de transacción, dado que los agentes informados arbitrarán la situación. Esto implica que las innovaciones del precio de error deben ser perturbaciones “pequeñas” en el precio de transacción. Basados en estos dos aspectos y teniendo el modelo de Glosten como punto de partida proponemos la siguiente especificación para los precios

$$\begin{aligned} m_t &= m_{t-1} + (1 - \theta)1(|\varepsilon_t| > r)\varepsilon_t, \\ s_t &= \theta 1(|\varepsilon_t| > r)\varepsilon_t + 1(|\varepsilon_t| \leq r)\varepsilon_t. \end{aligned}$$

El parámetro r recoge el coste de transacción no relacionado con el problema de se-

lección adversa (Glosten), y solo las perturbaciones mayores que este coste afectarán al precio eficiente m_t . Por otro lado, θ recogerá el coste debido a selección adversa. De esta especificación se obtiene la siguiente representación del precio de transacción

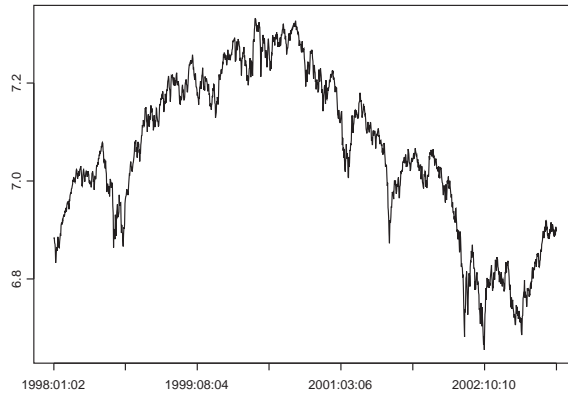
$$(1 - L)p_t = \begin{cases} \varepsilon_t - \theta\varepsilon_{t-1} & \text{si } |\varepsilon_t| > r \\ \varepsilon_t - \varepsilon_{t-1} & \text{si } |\varepsilon_t| \leq r. \end{cases} \quad (4.10)$$

El modelo (4.10) es un ARTIMA-Shock-Size. En este punto merece la pena puntualizar que aunque solo se observa una perturbación, ε_t , en el modelo estructural hay dos tipos de perturbaciones. Una relacionada con la nueva información, y por tanto, con el proceso de precios eficientes, y la otra originada en los agentes no informados y que será, por tanto, de naturaleza transitoria. Usando el razonamiento descrito anteriormente, estos dos tipos de perturbaciones estructurales se trasladan al proceso de los precios de transacción en términos de perturbaciones “grandes” y “pequeñas”. A través de los modelos ARTIMA somos capaces de contrastar la existencia de dicho mecanismo de revelación, y por lo tanto, de estos dos tipos de perturbaciones estructurales.

La Figura 4.2 representa el logaritmo de la serie diaria del *S&P500*. La Tabla 4.3 muestra la estimación de Mínimos Cuadrados del modelo lineal ARIMA y no lineal ARTIMA-Shock-Size ajustado a la serie diaria del *S&P500*. La estimación del modelo ARTIMA-Shock-Size fue obtenida numéricamente minimizando el cuadrado de los errores estimados usando el procedimiento de optimización DBCPOL de FORTRAN. Los estimadores de los parámetros de la media móvil en el modelo ARTIMA son claramente diferentes. Siguiendo el procedimiento de contraste descrito en la Sección 4.1, contrastamos en dos etapas la validez de este modelo. En la primera, contrastamos la hipótesis de linealidad frente a la alternativa de existencia de efecto threshold. Si rechazamos esta nula, en un segundo paso contrastaremos la nula de existencia de raíz unitaria en el régimen de las perturbaciones pequeñas.

Para contrastar linealidad, ajustamos un modelo ARIMA(0,1,1) al logaritmo de la serie,

Figura 4.2: Logaritmo del S&P500, 2 Enero 1998 - 29 Julio 2003



Cuadro 4.3: Estimación de Mínimos de Cuadrados de los modelos ARIMA(0,1,1) y ARTIMA(0,1,1,1) para los precios de los activos

Parámetros	θ	θ_1	θ_2	r	σ_ε	AIC
ARIMA(0, 1, 1)	0.0176 (0.0267)				0.0136	-5.7521
ARTIMA(0,1,1,1)		0.0131 (0.0268)	0.9717 (0.3632)	0.0035	0.0135	-5.7564

Nota: Entre paréntesis se encuentran las desviaciones estándar correspondientes.

construyendo la siguiente regresión auxiliar con los residuos estimados¹

$$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1} + \alpha_2 \hat{\varepsilon}_{t-1} 1(|\hat{\varepsilon}_{t-1}| < r) + u_t,$$

y contrastamos la significatividad de $\alpha_2(r)$. Dado que bajo la hipótesis nula el parámetro r no está identificado, el contraste propuesto es $\sup_r |t_{\alpha_2(r)}|$, donde $t_{\alpha_2(r)}$ es el t -ratio para la nula de $\alpha_2(r) = 0$. El resultado de este contraste se muestra en la Tabla 4.4.

La linealidad se rechaza al nivel de significación del 10 %, (p -valor =0.067). De la tabla 4.3 la hipótesis de raíz unitaria en el régimen MA correspondiente a las perturbaciones

¹Incluimos $\hat{\varepsilon}_{t-1}$ en la regresión auxiliar para detectar alguna posible asimetría en función del signo.

Cuadro 4.4: Contraste de Linealidad vs Efecto Threshold por el Tamaño de la Perturbación

Regresión Auxiliar					
$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1} + \alpha_2 \hat{\varepsilon}_{t-1} 1(\hat{\varepsilon}_{t-1} < r)$					
$\sup_r t_{\alpha_2} $	$p - value$	$\hat{\alpha}_0(\bar{r})$	$\hat{\alpha}_1(\bar{r})$	$\hat{\alpha}_2(\bar{r})$	\bar{r}
2.213	0.067	-0.000 (0.0003)	0.004 (0.0268)	-0.806 (0.3645)	0.0035

Nota: $\hat{\varepsilon}_t$ son los residuos del modelo lineal de la nula. El p-valor bootstrap es obtenido de 1000 replicaciones. En paréntesis las desviaciones estándar correspondientes.

pequeñas no puede ser rechazado a los niveles de significación habituales. Con estos resultados podemos concluir que los datos de *S&P500* analizados no rechazan la modelización ARTIMA-Shock-Size propuesta. Por lo tanto, para medir el tamaño del componente precio de error, s_t , obtenemos la descomposición B-N no lineal del modelo ARTIMA estimado. Para ello imponemos que $\theta_2 = 1$. Los resultados se muestran en la tabla 4.5. Nótese que como la teoría sugiere la contribución del componente paseo aleatorio es prácticamente la misma en todas las descomposiciones P-T, y el tamaño del componente transitorio satisface la desigualdad establecida en el Teorema 4.3.

Además de esta medida de la calidad del mercado de activos, el modelo ARTIMA-Shock-Size, provee información adicional. Por ejemplo, el valor estimado del parámetro threshold $\hat{r} = 0.00283$ (imponiendo $\theta_2 = 1$) implica que el coste de transacción estimado es 0.283 por ciento del precio del activo. Este valor es muy similar al obtenido por Hasbrouck (1993) para sus datos intradiarios. También, el porcentaje de perturbaciones puramente transitorias puede ser estimado, en este caso, en 23.65 %. Los periodos donde estas perturbaciones ocurren pueden ser identificados en este contexto como periodos de información no relevante.

Cuadro 4.5: Descomposición de la Varianza de las diferentes Descomposiciones $P - T$ para los Precios de los activos

	$B - N$ ARIMA(0,1,1)	$B - N$ ARTIMA(0,1,1,1)	$UC - 0$ ARIMA(0,1,1)
$\sigma_{\Delta m_t} \times 10^2$	1.338	1.370	1.339
$\sigma_{s_t} \times 10^2$	0.024	0.074	0.176

Nota: $p_t = m_t + s_t$, con m_t paseo aleatorio y s_t el componente transitorio. $\sigma_{\Delta m_t}$ y σ_{s_t} son la desviación estándar de Δm_t y s_t respectivamente.

4.5. Aplicación al PNB: Persistencia Asimétrica de las Perturbaciones a la Producción

El objetivo de esta sección es responder empíricamente una importante cuestión respecto a las fluctuaciones en la producción: ¿Es la persistencia de las perturbaciones en la producción asimétrica?. Esta cuestión ha sido analizada por Beaudry y Koop (1993), Hess e Iwata (1997) y Elwood (1998) entre otros. Todos estos autores tratan con las asimetrías producidas por el signo de la perturbación, y no hay una conclusión evidente. Los primeros autores dan una respuesta afirmativa respecto a la existencia de asimetrías, mientras que el resto contradice este resultado. Nosotros usaremos el modelo ARTIMA-Shock-Size propuesto en la tesis para concluir que las perturbaciones con tamaños diferentes tienen diferente efecto a largo plazo en la producción. Las perturbaciones pequeñas serán transitorias, mientras que las grandes serán permanentes.

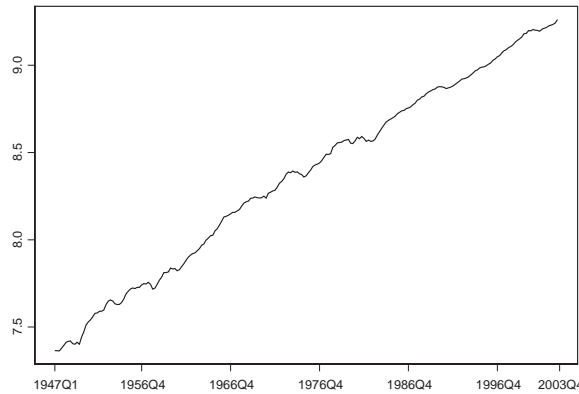
En esta sección, indirectamente, también trataremos de contestar la pregunta de qué tipo de perturbaciones es responsable de las fluctuaciones económicas. Esta cuestión ha sido analizada en detalle por Blanchard y Quah (1989), Cochrane (1994), y muchos otros autores. Cochrane (1994) pone de relieve dos aspectos fundamentales: (i) la identificación de las perturbaciones (ii) las perturbaciones endógenas. Sobre el último, se demandan nuevos mo-

delos teóricos, probablemente no lineales, que sean capaces de explicar las fluctuaciones del business cycle en términos de perturbaciones endógenas. Los modelos ARTIMA-Shock-Size son candidatos idóneos para tratar ambos aspectos simultáneamente.

La Figura 4.3 muestra el logaritmo del PNB real de EEUU para el periodo 1947:01-2003:03. Como en la aplicación anterior, la estimación del modelo ARTIMA-Shock-Size fue obtenida numéricamente minimizando el cuadrado de los errores estimados usando el procedimiento de optimización DBCPOL de FORTRAN. La Tabla 4.6 recoge los resultados de dicha estimación de Mínimos Cuadrados para diferentes modelos lineales ARIMA y no lineales ARTIMA-Shock-Size en la serie del PNB. Siguiendo a Campbell y Mankiw (1987), consideramos todos los modelos ARIMA con p y q menores que 3. Por cuestiones de espacio, la tabla 4.6 solo presenta los resultados para los modelos ARIMA(1,1,0), ARIMA(0,1,1) y ARIMA(1,1,1). Para el modelo ARTIMA-Shock-Size, el proceso de estimación seguido es el descrito en la Sección 3.3. El modelo ajustado a la serie es el siguiente ARTIMA(1,1,1,1),

$$(1 - \phi L)(1 - L)y_t = \mu + (1 - \theta_1 L)\varepsilon_t^L + (1 - \theta_2 L)\varepsilon_t^s.$$

Figura 4.3: Logaritmo del PNB Real, 1947:01-2003:03



La Tabla 4.6 muestra que todos los modelos lineales ajustan la serie de forma similar. Siguiendo el criterio AIC el mejor modelo lineal será un ARIMA(1,1,0), sin embargo, para

Cuadro 4.6: Estimación de Mínimos Cuadrados de los Modelos ARIMA y ARTIMA para PNB Real

Parámetros	ARIMA(1,1,0)	ARIMA(0,1,1)	ARIMA(1,1,1)	ARTIMA(1,1,1,1)
ϕ	0.335 (0.062)		0.453 (0.171)	0.647 (0.182)
θ		-0.260 (0.064)	0.131 (0.191)	
θ_1				0.320 (0.196)
θ_2				1.210 (0.396)
$r \times 10^2$				0.477
$\sigma_\varepsilon \times 10^2$	0.956	0.969	0.950	0.941
AIC	-6.452	-6.424	-6.447	-6.475

Nota: Entre paréntesis la desviación estándar correspondiente. Los resultados de la columna ARTIMA(1,1,1,1) corresponden a la estimación de la segunda etapa.

eliminar cierta correlación en los residuos, seleccionaremos el modelo ARIMA(1,1,1) para el contraste de linealidad. La estrategia de este contraste fue descrita en la Sección 4.2 y usada en la aplicación anterior. Para contrastar la nula de linealidad usamos la siguiente regresión auxiliar con los residuos del modelo lineal ARIMA (1,1,1)

$$\hat{\varepsilon}_t = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1} + \alpha_2 \hat{\varepsilon}_{t-1} 1(|\hat{\varepsilon}_{t-1}| < r) + u_t.$$

En esta regresión contrastaremos la significatividad de $\alpha_2(r)$ con el $\sup_r |t_{\alpha_2(r)}|$, donde $t_{\alpha_2(r)}$ es el t -ratio para la nula de $\alpha_2(r) = 0$. Los resultados se muestran en la Tabla 4.7. Claramente la nula de linealidad es rechazada ya que el p -valor es inferior a 0.05. Esto implica que el tamaño de la perturbación es capaz de generar asimetrías en su persistencia en el PNB.

Con la intención de contrastar si las perturbaciones pequeñas son transitorias o no, contrastamos que $\theta_2 = 1$. Del Teorema 3.5, el contraste se obtiene directamente a través

Cuadro 4.7: Contraste de Linealidad vs Efecto Threshold por el Tamaño de la Perturbación

Regresión Auxiliar					
$\widehat{\varepsilon}_t = \alpha_0 + \alpha_1 \widehat{\varepsilon}_{t-1} + \alpha_2 \widehat{\varepsilon}_{t-1} 1(\widehat{\varepsilon}_{t-1} < r)$					
$\sup_r t_{\alpha_2} $	$p - valor$	$\widehat{\alpha}_0(\bar{r})$	$\widehat{\alpha}_1(\bar{r})$	$\widehat{\alpha}_2(\bar{r})$	\bar{r}
2.362	0.045	0,000 (0,0006)	0,021 (0,0676)	-0,789 (0,3363)	0.004

Nota: $\widehat{\varepsilon}_t$ son los residuos bajo el modelo lineal. El p-valor bootstrap se obtiene de 1000 replicaciones. Entre paréntesis las desviaciones estándar correspondientes.

de la última columna de la Tabla 4.6, concluyendo que la nula de raíz unitaria en este régimen de la MA no puede ser rechazada. Por lo tanto, los datos del PNB considerados no rechazan que la perturbaciones grandes son permanentes mientras que las pequeñas serán transitorias.

Imponiendo $\theta_2 = 1$, calculamos la descomposición P-T ((4.4)-(4.5)) propuesta en la Sección 4.3. Ésta es

$$(1 - L) y_t^P = \mu + \frac{(1 - \theta_1 L)}{(1 - \phi L)} \varepsilon_t^L,$$

$$y_t^T = \frac{1}{(1 - \phi L)} \varepsilon_t^s,$$

donde ε_t^s corresponde a las perturbaciones menores que $\widehat{r} = 0.0046$. Esto implica que las perturbaciones menores que $\sigma_\varepsilon/2$, aproximadamente, serán transitorias. Estas representan el 50.7 % de las perturbaciones en el PNB (el 49.3 % de ellas serán permanentes).

Las Figuras 4.4, 4.5, y 4.6 representan gráficamente las descomposiciones P-T discutidas en la Sección 4.3, junto con los puntos de referencia para el business cycle del NBER. Numéricamente esta comparación se resume en la Tabla 4.8. Ambas descomposiciones, lineal y no lineal, de B-N producen resultados muy similares (las diferencias pueden deberse a cuestiones de estimación). Sin embargo, la descomposición no lineal P-T ortogonal es

claramente diferente. En primer lugar, el componente transitorio es menos volátil, y en segundo lugar, el componente permanente evoluciona de forma más suave.

Para finalizar, merece la pena comentar el nexo entre la perturbaciones grandes en el modelo ARTIMA (1,1,1,1) y los puntos de referencia del NBER para las crisis. En la Figura 4.7 se puede ver que los “picos” y “valles” del business cycles se encuentran relacionados con perturbaciones grandes de signo opuesto. Las situaciones en las que la producción abandona la crisis suelen ir acompañadas de perturbaciones grandes positivas, lo contrario ocurre cuando se abandona una época de crecimiento.

Figura 4.4: Descomposición B-N del PNB Real basada en el Modelo ARIMA(1,1,1)

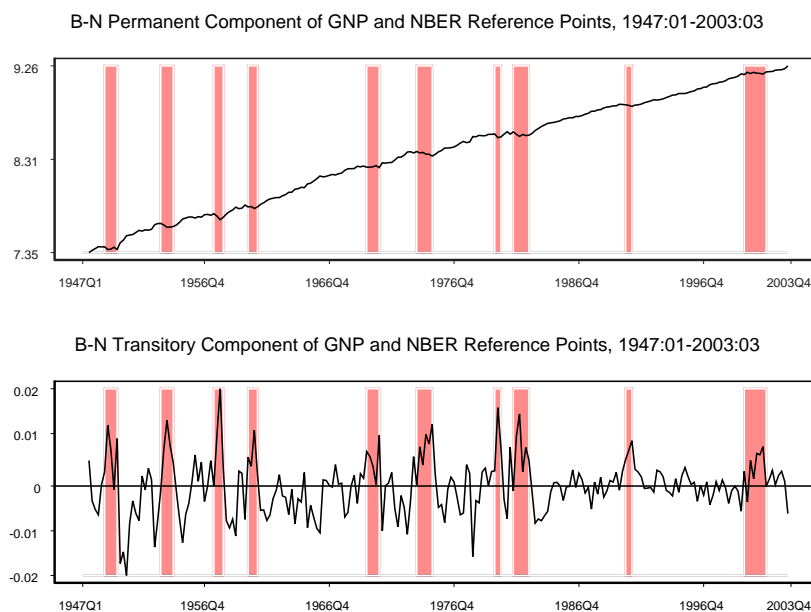


Figura 4.5: Descomposición B-N del PNB Real basada en el Modelo ARTIMA(1,1,1,1)

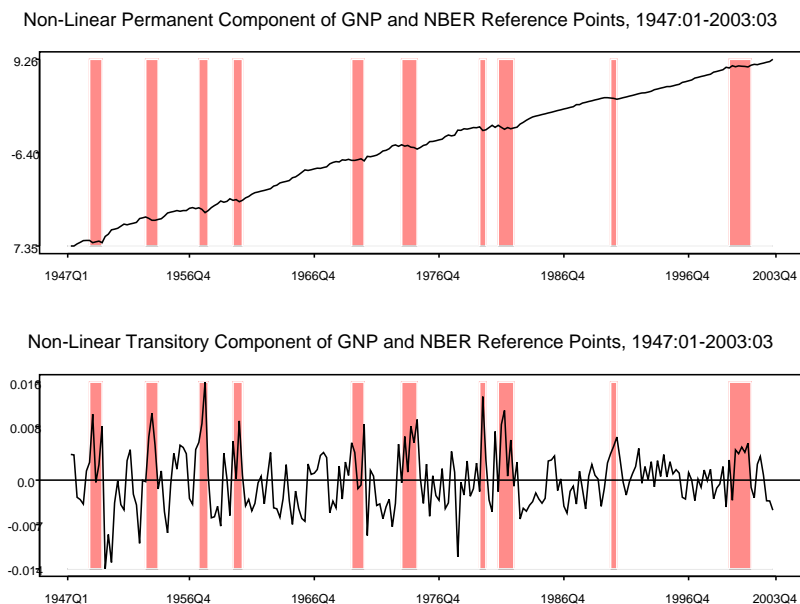
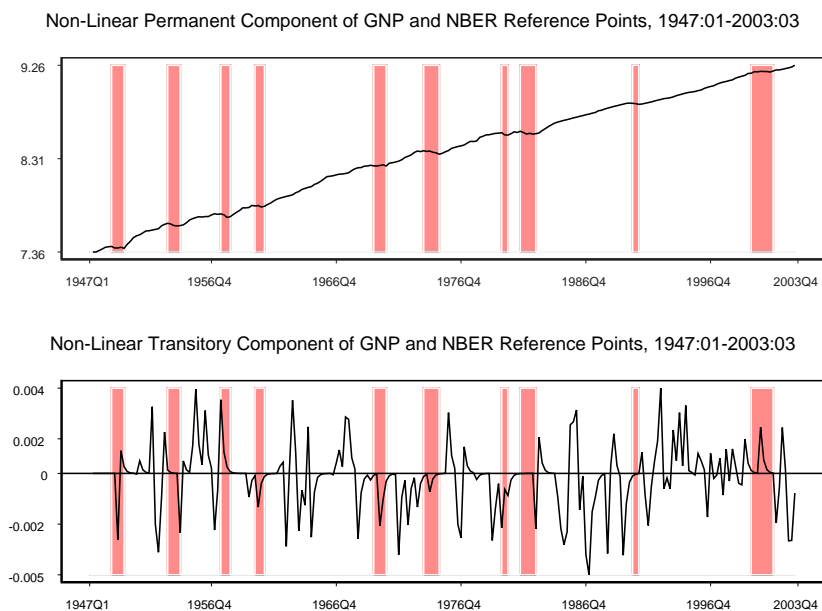


Figura 4.6: Descomposición Ortogonal del PNB Real basada en el Modelo ARTIMA(1,1,1,1)

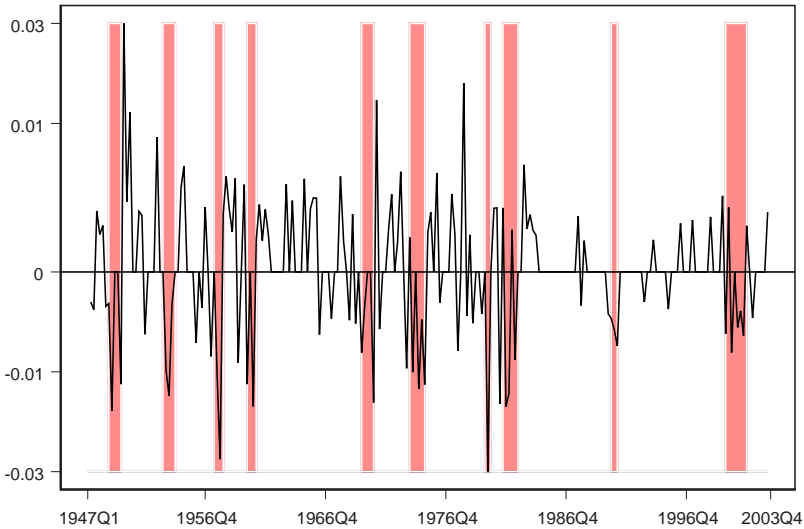


Cuadro 4.8: Varianza de las Diferentes descomposiciones P-T para el PNB Real

	$\sigma^2_{\Delta y_t^P} \times 10^3$	$\sigma^2_{y_t^T} \times 10^3$
B-N ARIMA(1,1,1)	0.2288	0.0398
B-N ARTIMA(1,1,1,1)	0.1825	0.0229
Ortogonal ARTIMA(1,1,1,1)	0.0982	0.0032

Nota: $y_t = y_t^P + y_t^T$. En la descomposición $B - N$ y_t^P es un paseo aleatorio. En la descomposición ortogonal ARTIMA y_t^P esta formada solo por las perturbaciones permanentes, esto es $E(\Delta y_t^P, y_t^T) = 0$ (ver (4.4) y (4.5)). $\sigma^2_{\Delta y_t^P}$ y $\sigma^2_{y_t^T}$ son las varianzas de Δy_t^P y y_t^T respectivamente.

Figura 4.7: Perturbaciones Grandes y Puntos de Referencia del NBER, 1947:01-2003:03



4.6. Conclusión

En este capítulo se lleva a cabo una aplicación, referente a la descomposición de una serie, de la capacidad de los modelos ARTIMA para identificar las perturbaciones de un sistema dinámico. Esta identificación se basa en el comportamiento a largo plazo de la perturbación (permanente o transitoria) y en su tamaño (grande o pequeña). Lo último es un caso especial más que una necesidad, ya que como hemos visto en el Capítulo 2, podría considerarse la característica de cualquier otra variable de interés. Los modelos ARTIMA presentan un marco adecuado para contrastar si una perturbación con una característica determinada (en este Capítulo su tamaño) tiene un efecto permanente o transitorio en la serie. Una vez que las perturbaciones son identificadas se construyen dos nuevas descomposiciones permanente-transitorio no lineales, que son comparadas con las descomposiciones permanente-transitorio lineales estándar $B - N$ y $UC - 0$ habituales en la literatura.

El capítulo considera dos aplicaciones. La primera estudia la calidad de los mercados de activos midiendo la importancia del componente transitorio (precio de error) de la serie de precios de activos. Esto se desarrolla siguiendo la propuesta de Hasbrouck (1993) y de acuerdo a una de las nuevas descomposiciones P-T no lineales presentadas en este capítulo. La segunda aplicación se realiza para el PNB y trata de contestar a la pregunta de si la persistencia de las perturbaciones a la producción es asimétrica o no. Elwood (1998) responde negativamente considerando como efecto diferenciador el signo de la perturbación, mientras que nuestra respuesta es claramente afirmativa cuando se considera el tamaño como la causa de la asimetría.

4.7. Bibliografía

Bai, J., (1994): “Weak convergence of the sequential empirical processes of residuals in ARMA models”. *The Annals of Statistics*, Vol. **22**, 2051-2061.

Brännäs, K., De Gooijer, J.G. y Teräsvirta, T. (1998): “Testing linearity against nonlinear moving average models”. *Commun. Statist. - Theory Meth.*, Vol. **27**, 2025-2035.

Beaudry, P. y Koop, G., (1993): “Do recessions permanently change output?”. *Journal of Monetary Economics*, Vol. **31**, 149-164.

Beveridge, S. y Nelson, C.R., (1981): “A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘Business Cycle’”. *Journal of Monetary Economics*, Vol. **7**, 151-174.

Blanchard, O. y Quah, D., (1989): “The dynamic effects of Aggregate Demand and Supply disturbances”. *The American Economic Review*, Vol. **79**, 655-673.

Brännäs, K., Gooijer, J.G. y Teräsvirta, T. (1998): “Testing linearity against nonlinear moving average models”. *Communications in Statistics, Theory and Methods*. Vol. **27**, pag. 2025-2035.

Campbell, J. y Mankiw, N.G., (1987): “Are Output Fluctuations Transitory?”. *Quarterly Journal of Economics*, Vol. **102**, 857- 880.

Clarida, R.H. y Taylor, M.P., (2003): “Nonlinear Permanent-Temporary Decompositions in Macroeconomics and Finance”. *The Economic Journal*, Vol. **113**, C125-C139.

Clark, P.K. y (1987): “The Cyclical Component of U.S. Economic Activity”. *Journal of Quarterly Journal of Economics*, Vol. **102**, 797-814.

Cochrane, J.H., (1994): “Shocks”. NBER working paper series; working paper no. 4698. Cambridge, MA.

Davis, R.B., (1977): “Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative”. *Biometrika*, Vol. **64**, 247-254.

Davies, R.B., (1987): “Hypothesis testing when a nuisance parameter is present only

under the alternative”. *Biometrika*, Vol. **74**, 33-43.

Elwood, S., (1998): “Is the Persistence of Shocks to Output Asymmetric?”. *Journal of Monetary Economics*, Vol. **41**, 411-426.

Engle, R.F. y Smith, A.D., (1999): “Stochastic Permanent Breaks”. *Review of Economics and Statistics*, Vol. **84**, 553.

Glosten, L.R., (1987): “Components of the Bid-Ask spread and the statistical properties of transaction prices”. *Journal of Finance*, Vol. **42**, 1293-1307.

Hamilton, J.D., (1989): “A new approach to the economic analysis of nonstationary time series and the business cycle”. *Econometrica*, Vol. **57**, 357-384.

Harvey, A.C., (1985): “Trends and Cycles in Macroeconomics Time Series”. *Journal of Business and Economic Statistics*, Vol. **3**, 216-27.

Hasbrouck, J., (1993): “Assessing the quality of a security market: A new approach to transaction-cost measurement”. *The Review of Financial Studies*, Vol. **6**, 191-212.

Hess, G.D. y Iwata, S., (1997): “Asymmetric persistence in GDP? A deeper look at depth”. *Journal of Monetary Economics*, Vol. **40**, 535-554.

Koul, H.L. y Stute, W., (1999): “Nonparametric model checks for time series”. *The Annals of Statistics*, Vol. **27**, 204-236.

Mortely, J.C., Nelson, C.R. y Zivot, E., (2001): “Why are Beveridge-Nelson and Unobserved-Component Decomposition of GDP so Different?”. *Review of Economics and Statistics*, Vol. **85**, 235-244.

Neftci, S.N., (1984): “Are Economic Time Series Asymmetric over the Business Cycle?”. *Journal of Political Economy*, Vol. **92**, 307-328.

Phillips, P.C.B. y Solo, V., (1992): “Asymptotics for Linear Processes”. *The Annals of Statistics*, Vol. **20**, 971-1001.

Pierce, D.A., (1979): “Signal Extraction Error in Nonstationary Time Series”. *The Annals of Statistics*, Vol. **7**, 1303-1320.

Quah, D., (1992): “The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds”. *Econometrica*, Vol. **60**, 107-118.

Roll, R., (1984): “A Simple model of the implicit Bid-Ask spread in an efficient market”. *Journal of Finance*, Vol. **39**, 1127-1446.

Watson, M.W., (1986): “Univariate detrending methods with stochastic trends”. *Journal of Monetary Economics*, Vol. **18**, 49-75.

Wold, H., (1938): *A Study in the Analysis of Stationary Time Series*. Uppsala: Almqvist & Wiksells.

4.8. Anexo

Prueba del Teorema 4.1

Por razones de simplicidad y sin pérdida de generalidad, este Teorema será probado para el caso $\mu = 0$, $p = 1$ y $q = 1$. Parte de la prueba hará uso de los resultados de Bai (1994).

Escribamos el proceso ARMA(1,1), representación de $x_t = (1 - L)y_t$, como $\varepsilon_t = x_t - \rho_0 x_{t-1} - \theta_0 \varepsilon_{t-1}$ y los residuos correspondientes como $\hat{\varepsilon}_t = x_t - \hat{\rho} x_{t-1} - \hat{\theta} \hat{\varepsilon}_{t-1}$, donde $(\hat{\rho}, \hat{\theta})$ son los estimadores LS de los parámetros (ρ_0, θ_0) . Estos estimadores satisfacen que $T^{-1/2}(\hat{\rho} - \rho_0)$ y $T^{-1/2}(\hat{\theta} - \theta_0)$ son $O_p(1)$.

Substrayendo los errores de los residuos, sustituyendo y haciendo uso de $\hat{\varepsilon}_0 = 0$, obtenemos que

$$\hat{\varepsilon}_t - \varepsilon_t = (-1)^{t-1} \hat{\theta}^t \varepsilon_0 - (\hat{\rho} - \rho_0) \sum_{j=0}^{t-1} (-1)^j \hat{\theta}^j x_{t-1-j} - (\hat{\theta} - \theta_0) \sum_{j=0}^{t-1} (-1)^j \hat{\theta}^j \varepsilon_{t-1-j}. \quad (4.11)$$

Denotamos $\hat{\phi} = (\hat{\theta}, \sqrt{T}(\hat{\rho} - \rho_0), \sqrt{T}(\hat{\theta} - \theta_0))$ y $\phi = (u, v, w) \in R^3$. Definimos

$$\Lambda_{\phi t} = u^t \varepsilon_0 + T^{-1/2} \left(v \sum_{j=0}^{t-1} u^j x_{t-1-j} + w \sum_{j=0}^{t-1} u^j \varepsilon_{t-1-j} \right) = u^t \varepsilon_0 + T^{-1/2} \xi_{\phi t}. \quad (4.12)$$

De (4.2) y la definición de $t_{\alpha(r)}$

$$t_{\alpha(r)} = \frac{\hat{\alpha}(r)}{\hat{\sigma}_{\hat{\alpha}}} = T^{-1/2} \sum_{t=2}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}^s(r) \left[T^{-1} \sum_{t=2}^T (\hat{\varepsilon}_{t-1}^s(r))^2 T^{-1} \sum_{t=1}^T \hat{u}_t^2(r) \right]^{-1/2}, \quad (4.13)$$

donde $\hat{\varepsilon}_t^s(r) = \hat{\varepsilon}_t 1(|\hat{\varepsilon}_t| < r)$ y $\hat{u}_t(r) = \hat{\varepsilon}_t - \hat{\alpha}(r) \hat{\varepsilon}_{t-1}^s(r)$.

La prueba consiste básicamente en descomponer cada término de (4.13) en distintas partes, y estudiar el comportamiento asintótico de cada una de esas partes. Para la primera

parte, $T^{-1/2} \sum_{t=2}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r)$, se prueba que,

$$T^{-1/2} \sum_{t=2}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r) = X_{1,T}(r) + X_{2,T}(r) + o_p(1),$$

donde

$$\begin{aligned} X_{1,T}(r) &\xrightarrow{d} D_1(r), \\ X_{2,T}(r) &\xrightarrow{d} D_2(r), \end{aligned}$$

uniformemente en r , con $D_1(r)$ y $D_2(r)$ distribuciones no degeneradas. Para la otra parte en (4.13), se prueba que,

$$T^{-1} \sum_{t=2}^T (\widehat{\varepsilon}_{t-1}^s(r))^2 T^{-1} \sum_{t=1}^T \widehat{u}_t^2(r) = X_{3,T}(r) + o_p(1),$$

donde

$$X_{3,T}(r) \xrightarrow{p} \Sigma(r),$$

uniformemente en r .

Comenzamos con $T^{-1/2} \sum_{t=2}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r)$, probaremos que es tight y que para todo $r \in [r, \bar{r}]$ converge a una distribución no degenerada. Definimos

$$\begin{aligned} \Upsilon_{x,t}(\widehat{\theta}) &= \sum_{j=0}^{t-1} (-1)^j \widehat{\theta}^j x_{t-1-j}, \\ \Upsilon_{\varepsilon,t}(\widehat{\theta}) &= \sum_{j=0}^{t-1} (-1)^j \widehat{\theta}^j \varepsilon_{t-1-j}. \end{aligned}$$

De (4.11), $T^{-1/2} \sum_{t=2}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r)$ se puede escribir como

$$\begin{aligned} T^{-1/2} \sum_{t=1}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_{t-1}^s(r) &= T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\widehat{\varepsilon}_{t-1}| < r) - \frac{(\widehat{\rho} - \rho_0)}{T^{1/2}} \sum_{t=1}^T \varepsilon_t \Upsilon_{x,t-1}(\widehat{\theta}) 1(|\widehat{\varepsilon}_{t-1}| < r) \\ &+ (\widehat{\theta} - \theta_0) T^{-1/2} \sum_{t=1}^T \varepsilon_t \Upsilon_{\varepsilon,t-1}(\widehat{\theta}) 1(|\widehat{\varepsilon}_{t-1}| < r) + (\widehat{\rho} - \rho_0) T^{-1/2} \sum_{t=1}^T \varepsilon_{t-1} \Upsilon_{x,t}(\widehat{\theta}) 1(|\widehat{\varepsilon}_{t-1}| < r) \\ &+ (\widehat{\theta} - \theta_0) T^{-1/2} \sum_{t=1}^T \varepsilon_{t-1} \Upsilon_{\varepsilon,t}(\widehat{\theta}) 1(|\widehat{\varepsilon}_{t-1}| < r) + R_{1,t}(\widehat{\rho}, \widehat{\theta}, r). \quad (4.14) \end{aligned}$$

Dado que

$$\begin{aligned} E \left[\sup_u \Upsilon_{x,t}^2(u) \right] &\leq K_{\Upsilon x} < \infty, \\ E \left[\sup_u \Upsilon_{\varepsilon,t}^2(u) \right] &\leq K_{\Upsilon \varepsilon} < \infty, \end{aligned}$$

es directo probar que

$$\sup_r R_{1,t}(\hat{\rho}, \hat{\theta}, r) = o_p(1).$$

El primer término de (4.14) puede ser escrito como

$$\begin{aligned} T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\hat{\varepsilon}_{t-1}| < r) &= T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r) \\ &\quad + T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} (1(|\hat{\varepsilon}_{t-1}| < r) - 1(|\varepsilon_{t-1}| < r)). \end{aligned} \quad (4.15)$$

Usando el Lema 3.1 de Koul y Stute (1999), se puede probar que el proceso empírico $T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r)$ es tight y,

$$T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\varepsilon_{t-1}| < r) \xrightarrow{d} D_1(r).$$

donde $D_1(r)$ es un proceso gaussiano con función de covarianzas dado por

$$M_1(r_1, r_2) = E(\varepsilon_t^2) E(\varepsilon_{t-1}^2 1(r_1 < |\varepsilon_{t-1}| < r_2))$$

para $r_2 > r_1$.

Para el último término de (4.15) necesitamos probar que

$$\sup_r T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} (1(|\hat{\varepsilon}_{t-1}| < r) - 1(|\varepsilon_{t-1}| < r)) = o_p(1).$$

Dado que $\varepsilon_t \varepsilon_{t-1} = \varepsilon_t^+ \varepsilon_{t-1}^+ + \varepsilon_t^+ \varepsilon_{t-1}^- + \varepsilon_t^- \varepsilon_{t-1}^+ + \varepsilon_t^- \varepsilon_{t-1}^-$, donde $\varepsilon_t^+ = \varepsilon_t 1(\varepsilon_t > 0)$ y $\varepsilon_t^- = \varepsilon_t 1(\varepsilon_t < 0)$, será suficiente con probar que

$$\sup_r \left| T^{-1/2} \sum_{t=1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ (1(\hat{\varepsilon}_{t-1} < r) - 1(\varepsilon_{t-1} < r)) \right| = o_p(1). \quad (4.16)$$

Para ello, usaremos la prueba del Teorema 1 en Bai (1994). Definimos

$$F(r) = E \left| \varepsilon_t^+ \right| E \left[\left| \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) \right| \right],$$

con $F(-\infty) = 0$ y $F(\infty) = K_F < \infty$. Entonces, usando (4.11) y (4.12)

$$\begin{aligned} T^{-1/2} \sum_{t=1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ (1(\widehat{\varepsilon}_{t-1} < r) - 1(\varepsilon_{t-1} < r)) &= \\ T^{-1/2} \sum_{t=1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ \left(1\left(\varepsilon_{t-1} < r + \Lambda_{\widehat{\phi}_{t-1}}\right) - 1(\varepsilon_{t-1} < r) \right) &= \\ = \widehat{K}_T(r) - K_T(r). \end{aligned}$$

Para estudiar este proceso, $\widehat{K}_T(r) - K_T(r)$, será suficiente con estudiar el proceso auxiliar,

$$G_T(x, \phi) = T^{-1/2} \sum_{t=1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ (1(\varepsilon_{t-1} < r + \Lambda_{\phi_{t-1}}) - 1(\varepsilon_{t-1} < r)).$$

La expresión (4.16) se obtiene del siguiente resultado

$$\sup_{\phi \in D_b} \sup_r |G_T(x, \phi)| = o_p(1) \quad \text{para todo } b > 0, \quad (4.17)$$

donde $D_b = [-\bar{\theta}, \bar{\theta}] \times [b, b]^2$ con $0 < \bar{\theta} < 1$ y $b > 0$. Definimos

$$\begin{aligned} Z_T(r, \phi) &= T^{-1/2} \sum_{t=1}^T \left(\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Lambda_{\phi_{t-1}}) - F(r + \Lambda_{\phi_{t-1}}) - \varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) + F(r) \right), \\ \Pi_T(r, \phi) &= T^{-1/2} \sum_{t=1}^T F(r + \Lambda_{\phi_{t-1}}) - F(r). \end{aligned}$$

Por lo tanto, $|G_T(x, \phi)| \leq |Z_T(r, \phi)| + |\Pi_T(r, \phi)|$. Para probar el resultado (4.17), será suficiente probar las siguientes dos proposiciones.

Proposición 4.4 *Bajo el supuesto A.0, se tiene*

$$\sup_{\phi \in D_b} \sup_r |Z_T(r, \phi)| = o_p(1) \quad \text{para todo } b > 0.$$

Proposición 4.5 *Bajo el supuesto A.0, se tiene*

$$\sup_{\phi \in D_b} \sup_r |\Pi_T(r, \phi)| = o_p(1) \quad \text{para todo } b > 0.$$

La prueba de ambas proposiciones sigue los mismos pasos que en Bai (1994) y por lo tanto ésta no se desarrollará en detalle, centrándonos únicamente en los aspectos relevantes. Nótese que $\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Lambda_{\phi_{t-1}})$ y $F(r)$ mantienen la propiedad de función creciente monótona en r de la función indicador de Bai.

Prueba de la Proposición 4.4. Define $\eta_{t-1} = C \sum_{j=1}^{t-1} \tau^j (|x_{t-j-1}| + |\varepsilon_{t-j-1}|)$ para algún $C > 0$ y $\tau \in (0, 1)$, y para todo $\lambda \in R$,

$$\begin{aligned} \tilde{Z}_T(r, \phi, \lambda) &= T^{-1/2} \sum_{t=1}^T (\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Gamma_{t-1}(\phi, \lambda)) - F(r + \Gamma_{t-1}(\phi, \lambda))) \\ &\quad - T^{-1/2} \sum_{t=1}^T (\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) + F(r)), \end{aligned}$$

donde $\Gamma_t(\phi, \lambda) = u^t \varepsilon_0 + \lambda t \bar{\theta}^{T-1} |\varepsilon_0| + T^{-1/2} \xi_{\phi t} + \lambda T^{-1/2} \eta_t$. Nótese que $\tilde{Z}_T(r, \phi, 0) = Z_T(r, \phi)$.

Usando que D_b es compacto, este conjunto puede ser particionado en un número finito de subconjuntos tal que el diámetro de cada subconjunto sea menor que ζ . Denote este subconjunto por $\Delta_1, \Delta_2, \dots, \Delta_{m(\zeta)}$. Fijando k y considerando Δ_k y $\phi_k = (u_k, v_k, w_k) \in \Delta_k$. Procediendo como en Bai (1994), elegimos C suficientemente grande y $\tau \in (0, 1)$ tal que

$$|\Lambda_{\phi_k t} - \Lambda_{\phi t}| \leq \zeta t \bar{\theta}^{T-1} |\varepsilon_0| + \zeta T^{-1/2} \eta_t \quad \text{para todo } \phi \in \Delta_k.$$

Con esta desigualdad y la monotoneidad de las funciones de $Z_T(r, \phi)$, se puede probar que

$$Z_T(r, \phi) \leq \tilde{Z}_T(r, \phi_k, \zeta) + T^{-1/2} \sum_{t=1}^T F(r + \Gamma_{t-1}(\phi_k, \zeta)) - F(r + \Lambda_{\phi_{t-1}}),$$

para todo $\phi \in \Delta_k$, y la igualdad inversa con ζ sustituida por $-\zeta$. Ahora será necesario el siguiente lema, el cual se prueba en Bai (1994).

Lema 4.1 *Bajo el supuesto A.0, para todo $\phi = (u, v, w) \in D_b$ y todo $\lambda \in R$ dado, tenemos que*

$$(a) \quad T^{-1/2} \sum_{t=1}^T \left(|u^t \varepsilon_0| + t \bar{\theta}^{t-1} |\lambda \varepsilon_0| \right) = o_p(1),$$

$$(b) \quad T^{-1/2} \max_{1 \leq t \leq T} (|\xi_{\phi t}| + |\lambda \eta_t|) = o_p(1),$$

$$(c) \quad T^{-1} \sum_{t=1}^T (|\xi_{\phi t}| + |\lambda \eta_t|) = O_p(1).$$

Usando este lema y la siguiente desigualdad

$$T^{-1/2} \left| \sum_{t=1}^T F(r + \Gamma_{t-1}(\phi_k, \zeta)) - F(r + \Lambda_{\phi t-1}) \right| \leq \frac{2\zeta f_e^*}{T^{1/2}} \sum_{t=1}^T \left((t-1) \bar{\theta}^{t-2} |\varepsilon_0| + \frac{\eta_{t-1}}{T^{1/2}} \right) = \zeta O_p(1),$$

donde $f_e^* = \max_e ef(e) < K_{fe} < \infty$, es directo probar que

$$\sup_{\phi \in D_b} \sup_r T^{-1/2} \left| \sum_{t=1}^T F(r + \Gamma_{t-1}(\phi_k, \zeta)) - F(r + \Lambda_{\phi t-1}) \right| = \zeta O_p(1).$$

El $O_p(1)$ es uniforme en r y ϕ . Por lo tanto,

$$\sup_{\phi \in D_b} \sup_r |Z_T(r, \phi)| \leq \max_{k \leq m(\zeta)} \sup_r \left| \tilde{Z}_T(r, \phi_k, \zeta) \right| + \max_{k \leq m(\zeta)} \sup_r \left| \tilde{Z}_T(r, \phi_k, -\zeta) \right| + \zeta O_p(1).$$

Entonces, $\zeta O_p(1)$ puede ser tan pequeño como queramos en probabilidad. La proposición 4.4 se obtiene si los dos primeros términos son $o_p(1)$. Esto será cierto si

$$\sup_r \left| \tilde{Z}_T(r, \phi, \zeta) \right| = o_p(1) \quad \text{para todo } \phi, \zeta. \quad (4.18)$$

Para ello será necesario el siguiente lema.

Lema 4.2 Para todo $\ell \in (0, 1/2)$, todo $\phi = (u, v, w) \in D_b$ y todo $\lambda \in R$,

$$\sup_{(r_1, r_2) \in B_{T, \ell}} T^{-1/2} \sum_{t=1}^T |F(r_2 + \Gamma_t(\phi, \lambda)) - F(r_1 + \Gamma_t(\phi, \lambda))| = o_p(1),$$

donde $B_{T, \ell} = \{(r_1, r_2) \in R \times R; |F(r_1) - F(r_2)| \leq K_{F,1} T^{-1/2-\ell}\}$ para algún $K_{F,1} < \infty$.

Dada la continuidad uniforme de $F(r)$ y que es acotada, la prueba de este lema es análoga a la del lema 2 de Bai (1994).

Ahora dividimos la recta real en $N(T) = \lceil T^{-1/2-\ell} \rceil + 1$ partes, con $\ell \in (0, 1/2)$, por los puntos $0 = r_0 < r_1 < \dots < r_{N(T)} = \infty$ con $F(r_i) = iN(T)^{-1}K_{F,1}$. Además con $r_k < r < r_{k+1}$, dado $\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r)$ y $F(r)$ son no decrecientes, tenemos que

$$\begin{aligned} \left| \tilde{Z}_T(r, \phi, \zeta) \right| &\leq \left| \tilde{Z}_T(r_{k+1}, \phi, \zeta) \right| + T^{-1/2} \sum_{t=1}^T [F(r_k + \Gamma_{t-1}) - F(r + \Gamma_{t-1})] \\ &\quad + T^{-1/2} \sum_{t=1}^T [\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r_{k+1}) - F(r_{k+1}) - \varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) + F(r)], \end{aligned}$$

y la igualdad inversa con r_{k+1} sustituida por r_k . Por lo tanto,

$$\begin{aligned} \sup_r \left| \tilde{Z}_T(r, \phi, \zeta) \right| &\leq \max_k \left| \tilde{Z}_T(r_k, \phi, \zeta) \right| \\ &\quad + \max_k T^{-1/2} \left| \sum_{t=1}^T F(r_{k+1} + \Gamma_{t-1}) - F(r_k + \Gamma_{t-1}) \right| \quad (4.19) \\ + \sup_{|g-g'| \leq K_F N(T)^{-1}} T^{-1/2} \sum_{t=1}^T &(\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g)) - g - \varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g')) + g') . \quad (4.20) \end{aligned}$$

Para la expresión (4.20), es necesario probar que $T^{-1/2} \sum_{t=1}^T (\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g)) - g)$ es tight. Este proceso puede ser escrito como

$$\begin{aligned} T^{-1/2} \sum_{t=1}^T (\varepsilon_t^+ \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g)) - g) &= T^{-1/2} \sum_{t=1}^T (\varepsilon_t^+ - E(\varepsilon_t^+)) \varepsilon_{t-1}^+ [1(\varepsilon_{t-1} < F^{-1}(g))] \\ &\quad + E(\varepsilon_t^+) T^{-1/2} \sum_{t=1}^T [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < F^{-1}(g)) - g/E(\varepsilon_t^+)] . \end{aligned}$$

La tightness de ambos procesos puede probarse usando el lema 3.1 de Koul y Stute (1999), y por lo tanto, se omite. El $o_p(1)$ de la expresión (4.20) se sigue de la tightness del proceso y del hecho que $K_{F,1}N(T)^{-1} = o_p(1)$. El $o_p(1)$ de la expresión (4.19) se sigue del lema 4.2.

Finalmente, $\tilde{Z}_T(r, \phi_k, \zeta)$ puede ser expresado como,

$$\begin{aligned} \tilde{Z}_T(r, \phi_k, \zeta) &= T^{-1/2} \sum_{t=1}^T [(\varepsilon_t^+ - E(\varepsilon_t^+)) (\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Gamma_{t-1}) - \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r))] \\ &+ T^{-1/2} E(\varepsilon_t^+) \sum_{t=1}^T [\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r + \Gamma_{t-1}) - G(r + \Gamma_{t-1}) - \varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r) + -G(r)], \end{aligned}$$

con $\Gamma_{t-1} = \Gamma_{t-1}(\phi, \zeta)$ y $G(r) = E[\varepsilon_{t-1}^+ 1(\varepsilon_{t-1} < r)]$. Entonces, los procesos de ambas sumas son secuencias de martingalas, y podemos aplicar Bai (1994) para obtener que

$$\max_k \left| \tilde{Z}_T(r_k, \phi, \delta) \right| = o_p(1)$$

■

Prueba de la Proposición 4.5. Esta prueba es análoga a la de la proposición 2 de Bai, y por tanto se omite. ■

Con esto, se prueba el resultado de (4.16). Extendiendo este resultado al resto de términos, se prueba que

$$\sup_r T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} (1(|\hat{\varepsilon}_{t-1}| < r) - 1(|\varepsilon_{t-1}| < r)) = o_p(1).$$

Por lo tanto, volviendo a la expresión (4.15), hemos probado que

$$T^{-1/2} \sum_{t=1}^T \varepsilon_t \varepsilon_{t-1} 1(|\hat{\varepsilon}_{t-1}| < r) \xrightarrow{d} D_1(r) + o_p(1),$$

donde el $o_p(1)$ es uniforme en r .

Ahora, para el resto de términos en (4.14), de forma similar se puede probar que

$$\begin{aligned} T^{-1} \sum_{t=1}^T \varepsilon_t \Upsilon_{x,t-1}(\hat{\theta}) 1(|\hat{\varepsilon}_{t-1}| < r) &= o_p(1) \\ T^{-1} \sum_{t=1}^T \varepsilon_t \Upsilon_{\varepsilon,t-1}(\hat{\theta}) 1(|\hat{\varepsilon}_{t-1}| < r) &= o_p(1) \\ T^{-1} \sum_{t=1}^T \varepsilon_{t-1} \Upsilon_{x,t}(\hat{\theta}) 1(|\hat{\varepsilon}_{t-1}| < r) &= T^{-1} \sum_{t=1}^T \varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| < r) + o_p(1) \\ T^{-1} \sum_{t=1}^T \varepsilon_{t-1} \Upsilon_{\varepsilon,t}(\hat{\theta}) 1(|\hat{\varepsilon}_{t-1}| < r) &= T^{-1} \sum_{t=1}^T \varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| < r) + o_p(1), \end{aligned}$$

donde el $o_p(1)$ es uniforme en r . Ahora, dado que $T^{1/2}[(\hat{\theta} - \theta_0), (\hat{\rho} - \rho_0)]$ converge a (V_θ, V_ρ) las cuales son distribuciones normales multivariantes de media 0 y matriz de varianzas y covarianzas $M_{\theta, \rho}$, es directo probar, usando los resultados anteriores y Slutsky, que

$$T^{-1/2} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-1}^s(r) \xrightarrow{d} D_1(r) + (V_\theta + V_\rho) M_s(r), \quad (4.21)$$

con $\Sigma_\varepsilon = E(\varepsilon_{t-1}^2 1(|\varepsilon_{t-1}| < r))$, $Cov(D_1(r), V_\phi) = E(\varepsilon_t^2) E(\varepsilon_{t-1} s_{1,t-1}(\lambda^0)) H_{1,1}^{-1}(\lambda^0)$ y $Cov(D_1(r), V_\theta) = E(\varepsilon_t^2) E(\varepsilon_{t-1} s_{2,t-1}(\lambda^0)) H_{2,2}^{-1}(\lambda^0)$.

Finalmente, es necesario probar que

$$T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_{t-1}^s(r))^2 \xrightarrow{p} \Sigma_\varepsilon(r), \quad (4.22)$$

$$T^{-1} \sum_{t=1}^T \hat{u}_t^2(r) \xrightarrow{p} \sigma_\varepsilon^2, \quad (4.23)$$

uniformemente en $r \in [\underline{r}, \bar{r}]$ con $0 < \Sigma_\varepsilon(r) < \infty$ y $0 < \Sigma_u(r) < \infty$.

Las pruebas de (4.22) y (4.23) son tediosas y similares a las pruebas anteriores. Por lo tanto, solo presentaremos su esquema. Para (4.22), nótese que $(\hat{\varepsilon}_t^s(r))^2$ puede ser escrito como

$$\begin{aligned} (\hat{\varepsilon}_t^s(r))^2 &= [\varepsilon_t + (\varepsilon_t - \hat{\varepsilon}_t)]^2 1(|\hat{\varepsilon}_t| < r) \\ &= \varepsilon_t^2 1(|\varepsilon_t| < r) + \varepsilon_t^2 [1(|\hat{\varepsilon}_t| < r) - 1(|\varepsilon_t| < r)] + \\ &\quad (\varepsilon_t - \hat{\varepsilon}_t)^2 1(|\hat{\varepsilon}_t| < r) + 2\varepsilon_t (\varepsilon_t - \hat{\varepsilon}_t) 1(|\hat{\varepsilon}_t| < r). \end{aligned}$$

Sobre el primer término, $T^{-1} \sum_{t=1}^T \varepsilon_t^2 1(|\varepsilon_t| < r)$, dado el supuesto A.0, es directo probar,

$$T^{-1} \sum_{t=1}^T \varepsilon_t^2 1(|\varepsilon_t| < r) \xrightarrow{p} \Sigma_\varepsilon(r),$$

uniformemente en $r \in [\underline{r}, \bar{r}]$ con $0 < \Sigma_\varepsilon(r) < \infty$. El resto de términos se pueden tratar como $T^{-1} \sum_{t=1}^T \varepsilon_t^+ \varepsilon_{t-1}^+ (1(\hat{\varepsilon}_{t-1} < r) - 1(\varepsilon_{t-1} < r))$, $R_{1,t}(\hat{\rho}, \hat{\theta}, r)$ y

$T^{-1} \sum_{t=1}^T \varepsilon_t \Upsilon_{x,t-1} \left(\widehat{\theta} \right) 1(|\varepsilon_{t-1}| < r)$ respectivamente. Una vez que (4.22) es probado, será directo probar que

$$\widehat{\alpha}(r) \xrightarrow{p} 0, \quad (4.24)$$

uniformemente en r . De (4.22) y (4.24), es fácil obtener (4.23).

El Teorema 4.1,

$$\sup_r t_{\alpha(r)}(r) \xrightarrow{d} \sup_r D(r),$$

con

$$D(r) \equiv [D_1(r) + (V_\theta + V_\rho)\Sigma_\varepsilon(r)] [\Sigma_\varepsilon(r) \sigma_\varepsilon^2]^{-1/2},$$

se sigue directamente de (4.21), (4.22), y (4.23).

Prueba del Teorema 4.3

Por simplicidad y sin perdida de generalidad presentamos la prueba para el caso de $\theta_2 = 1$.

Sea $\sigma_x^2 = V(x_t)$, y $\rho_{x1} = Cov(x_t, x_{t-1})$ con $x_t = (1 - L)y_t$. Si y_t es un ARTIMA(0,1,1,1), entonces,

$$\begin{aligned} \sigma_x^2 &= 2\sigma_\varepsilon^2 - (1 - \theta_1^2)\sigma_{\varepsilon,L}^2, \\ \rho_{x1} &= -\sigma_\varepsilon^2 + (1 - \theta_1)\sigma_{\varepsilon,L}^2, \end{aligned} \quad (4.25)$$

donde $\sigma_{\varepsilon,L}^2 = E[(\varepsilon_t^L)^2]$. De las ecuaciones (4.8) y (4.9),

$$\begin{aligned} V^{ARTIMA}(\Delta y_t^P) &= \sigma_x^2 + 2\rho_{x1}, \\ V^{ARTIMA}(y_t^T) &= -\rho_{x1} + (\theta_1^2 - \theta_1)\sigma_\varepsilon^2. \end{aligned}$$

Para obtener los correspondientes resultados para la descomposición $B - N$ lineal, usamos la representación de Wold de los procesos ARTIMA(0,1,1,1),

$$(1 - L)y_t = \mu + u_t - \theta u_{t-1}, \quad (4.26)$$

con u_t ruido blanco, ($E(u_t) = 0$, $\sigma_u^2 = E(u_t^2)$). Entonces,

$$\begin{aligned} V^{BN}(\Delta y_t^P) &= \sigma_x^2 + 2\rho_{x1}, \\ V^{BN}(y_t^T) &= -\theta\rho_{x1}. \end{aligned}$$

Finalmente, de forma similar, la descomposición $UC - 0$ de (4.26) es

$$\begin{aligned} (1 - L)y_t^P &= \eta_t^P, \\ y_t^T &= \eta_t^T, \end{aligned}$$

con $E[(\eta_t^P)^2] = \sigma_{\eta^P}^2$, $E[(\eta_t^T)^2] = \sigma_{\eta^T}^2$, $E(\eta_t^P, \eta_t^T) = 0$, y tal que $\sigma_{\eta^P}^2 = \sigma_x^2 + 2\rho_{x1}$ y $\sigma_{\eta^T}^2 = -\rho_{x1}$.

Por lo tanto,

$$\begin{aligned} V^{UC-0}(\Delta y_t^P) &= \sigma_x^2 + 2\rho_{x1}, \\ V^{UC-0}(y_t^T) &= -\rho_{x1}. \end{aligned}$$

Entonces, la $Var(\Delta y_t^P)$ es la misma para las tres descomposiciones. Con respecto a $Var(y_t^T)$, usando el hecho de que los parámetros θ y θ_1 son menores que 1, es directo probar que el máximo se obtiene para la descomposición $UC - 0$. Para probar que $V^{BN}(y_t^T)$ es menor o igual que $V^{ARTIMA}(y_t^T)$, es suficiente con probar que $\sigma_u^2 \geq \sigma_\varepsilon^2$. Esto se sigue de las siguientes dos expresiones,

$$\begin{aligned} V^{BN}(y_t^T) &= \sigma_x^2 - \sigma_u^2, \\ V^{ARTIMA}(y_t^T) &= \sigma_x^2 - \sigma_\varepsilon^2. \end{aligned}$$

De la representación de Wold (4.26),

$$\sigma_u^2 = \frac{\sigma_x^2 \pm \sqrt{\sigma_x^4 - 4\rho_{x1}^2}}{2}.$$

Dado que $\rho_{x1} = -\theta\sigma_u^2$ y $|\theta| < 1$, la solución válida es,

$$\sigma_u^2 = \frac{\sigma_x^2 + \sqrt{\sigma_x^4 - 4\rho_{x1}^2}}{2}.$$

Sustituyendo σ_x^2 y ρ_{x1} en términos de θ_1 , σ_ε^2 y $\sigma_{\varepsilon,L}^2$ de (4.25), tenemos

$$\sigma_x^4 - 4(\rho_{x1})^2 = \sigma_{\varepsilon,L}^4(1 - \theta_1^2)^2 + (\sigma_\varepsilon^2\sigma_{\varepsilon,L}^2 - \sigma_{\varepsilon,L}^4)(2 - 2\theta_1)^2.$$

Entonces, usando que $\sigma_x^2 = 2\sigma_\varepsilon^2 - (1 - \theta_1^2)\sigma_{\varepsilon,L}^2$ y $\sigma_\varepsilon^2 > \sigma_{\varepsilon,L}^2$ se obtiene que $\sigma_u^2 \geq \sigma_\varepsilon^2$ y por lo tanto el Teorema.

Capítulo 5

Modelos Threshold y Larga Memoria

Resumen: *En este Capítulo analizamos la relación de los modelos threshold con las propiedades que presentan los modelos de larga memoria. Este análisis se engloba dentro de la reciente literatura que relaciona los modelos con cambios estructurales y los de larga memoria. Para ello, nos centraremos en los modelos threshold autoregresivos con raíz unitaria, modelos TUR, así como en los modelos threshold introducidos a lo largo de esta tesis, los modelos ARTIMA, no analizados hasta ahora en este contexto. En concreto, veremos las condiciones necesarias para que este tipo de modelos generen ciertas propiedades utilizadas para definir los procesos con larga memoria. Igualmente, a través de un estudio de Monte Carlo, analizaremos los efectos que este tipo de modelos tienen en los contrastes de larga memoria y raíces unitarias.*

5.1. Introducción

La persistencia es una de las características más importantes de las series temporales, no solo por su relevancia teórica, sino también por las consecuencias que tiene en la teoría estadística. En un principio, la persistencia de las series se caracterizó a través de los procesos $I(0)$, o no integrados, y los procesos $I(1)$ o integrados de orden 1. Sin embargo, no siempre es posible aceptar la hipótesis nula de un tipo de proceso frente al otro de forma uniequívoca. Así, los procesos fraccionalmente integrados, $FI(d)$, cubrieron el hueco existente entre los procesos $I(0)$ e $I(1)$, abriendo una posibilidad a los casos en que ninguno de los dos tipos de procesos parecen adecuados. Estos casos, $0 < d < 1$, son los que denominaremos larga memoria a lo largo de la tesis. Sin embargo, los procesos $FI(d)$ implican una dependencia de los infinitos retardos de la serie difícil de justificar a través de modelos económicos teóricos. Esto ha echo que se haya desarrollado una amplia literatura que trata de explicar y justificar la existencia de este tipo de procesos a través de modelos más intuitivos desde el punto de vista económico.

Uno de los primeros trabajos en esta dirección es el de Granger (1980). Granger prueba que bajo determinadas condiciones la agregación sectorial de procesos $ARMA(p,q)$ puede generar larga memoria. Para ello, será necesaria cierta heterogeneidad en los parámetros de las series $ARMA(p,q)$ que se agregan, que deberán seguir una determinada distribución Beta. Posteriormente, Lippi y Zaffaroni (1999) sustituyen este supuesto de Granger sobre la distribución por un supuesto semiparamétrico más débil. Por otro lado, Chambers (1998) considera la agregación temporal, y no sólo sectorial. Otro método de generar larga memoria, también relacionado con la agregación, ha sido estudiado por Cioczek-Georges y Mandelbrot (1995) y Taqqu *et al.* (1997). En este caso se agregan variables binarias. Cada variable toma el mismo valor durante un tiempo n que es aleatorio. Si la función de distribución de n presenta colas pesadas, los autores prueban que las series agregadas, convenientemente estandarizadas, convergen a un movimiento Browniano Fraccional (MBF).

Relacionada con esta aproximación se encuentran los modelos de duración del error de Parke (1999). En estos modelos la serie está formada por la agregación de perturbaciones, de forma que cada perturbación dura un tiempo aleatorio n , que igualmente debe seguir una distribución de cola pesada. En este caso, no es necesaria la agregación para obtener ciertas propiedades de larga memoria, aunque como prueban Davidson y Sibbertsen (2004), no convergerán a un MBF.

Otro tipo de aproximación se consigue a través de los cambios estructurales y no linealidades. Así, Granger y Teräsvirta (1999) consideran el modelo no lineal:

$$y_t = \text{sign}(y_{t-1}) + \varepsilon_t$$

y muestran que puede generar larga memoria a medida que la varianza de ε_t disminuye, aunque no prueban una relación teórica. Posteriormente, Granger y Hyung (1999) generan propiedades de larga memoria a través del modelo media más ruido. Para ello es necesario que la probabilidad de cambio en la media dependa del tamaño muestral. En la misma dirección, Diebold e Inoue (2001) obtienen similares resultados para los modelos media más ruido, STOPBREAK de Engle y Smith (1999) y modelos de cambio de régimen de Markov. Todos estos casos se caracterizan porque algún parámetro del modelo depende del tamaño muestral T , bien la probabilidad de cambio estructural o alguno relacionado con la no linealidad del modelo. Si estos parámetros son constantes, ninguno de estos modelos generarán propiedades de larga memoria. Finalmente, Gouriéroux y Jasiak (2001) usan las dos aproximaciones, colas pesadas y probabilidad dependiente del tamaño muestral, para modelos de breaks infrecuentes.

Si entendemos la imposibilidad de aceptar de forma uniequívoca la H_0 de un proceso $I(0)$ o $I(1)$ frente al otro, como la imposibilidad de aceptar que todas las perturbaciones son o transitorias o permanentes, los modelos ARTIMA suponen una aproximación diferente a este problema. Los procesos $FI(d)$ proponen una solución que implica que todas las perturbaciones presentan un efecto intermedio entre el permanente de los $I(1)$ y el transitorio de

los $I(0)$. Los procesos ARTIMA, por el contrario, permiten un enfoque diferente en el que ni todas las perturbaciones son permanentes (caso $I(1)$) ni todas transitorias (caso $I(0)$), sino que unas son transitorias y otras permanentes dependiendo de un criterio económico que se determina a través de la variable threshold.

En este capítulo veremos las condiciones bajo las cuales los modelos Threshold son capaces de generar alguna propiedad de larga memoria de los procesos $FI(d)$, tanto a través de la agregación como de parámetros dependientes del tamaño muestral. Igualmente se realiza un detallado estudio de Monte Carlo que corrobora estos resultados teóricos y que servirá para analizar los efectos que los modelos threshold estudiados tienen sobre los contrastes de larga memoria, incluidos los de raíces unitarias.

El capítulo continua como sigue. En la Sección 2, se prueban las principales propiedades de los modelos de larga memoria que sirven para definirla. En la Sección 3, se obtienen los resultados teóricos que prueban la relación entre los modelos threshold y la larga memoria. En la Sección 4 se realiza el estudio de Monte Carlo. Finalmente, la Sección 5 presenta las conclusiones.

5.2. Definiciones de Larga Memoria

En esta sección realizaremos un breve resumen de las distintas definiciones de larga memoria, consideradas en la literatura, y sus relaciones. Cabe destacar que a lo largo de esta sección no trataremos la definición de larga memoria centrada en los casos de $I(1)$ o raíz unitaria, sino en los casos de procesos integrados fraccionales $FI(d)$, con $0 < d < 1$.

Tradicionalmente, la larga memoria en el dominio temporal se ha definido a través de la tasa de decaimiento de la función de autocorrelación cuando el retardo tiende hacia infinito. En concreto, para un proceso estacionario en covarianza, x_t , la definición de larga

memoria vendrá dada por:

$$\gamma_x(k) \simeq ck^{2d-1} \quad k \rightarrow \infty$$

con $\gamma_x(k)$ la función de autocorrelación de x_t . Para $d > 0$ la función de autocorrelación de un proceso de larga memoria se caracteriza por un decaimiento hiperbólico, más lento que el exponencial característico de los procesos $I(0)$. En el dominio de frecuencias, la definición se basará en las frecuencias bajas del espectro:

$$f_x(\varpi) \simeq g\varpi^{-2d} \quad \varpi \rightarrow 0^+$$

con f_x la densidad espectral de x_t . Ambas definiciones coincidirán bajo determinadas condiciones, dadas, por ejemplo, en Beran (1994). Una propiedad de los procesos de larga memoria que se obtiene como consecuencia de las definiciones anteriores, se basa en la tasa de convergencia de las varianzas de las sumas parciales del proceso x_t , $S_T = \sum_{t=1}^T x_t$, concretamente,

$$Var(S_T) = O(T^{2d+1})$$

Para relacionar esta propiedad con el espectro, se ha de tener en cuenta que el límite del espectro cuando la frecuencia ϖ tiende a cero es el límite de $T^{-1}S_T$.

5.3. Modelos Threshold y Larga Memoria

En esta sección analizaremos bajo que condiciones los modelos threshold pueden generar alguna de las propiedades o definiciones de larga memoria descritas en la sección anterior. Para ello consideramos dos tipos de modelos threshold. Por un lado, los modelos Threshold Unit Root (TUR), introducidos por González y Gonzalo (1998), y que en su forma más simple tienen la siguiente expresión,

$$y_t = \begin{cases} y_{t-1} + \varepsilon_t & z_t \in A \\ \rho y_{t-1} + \varepsilon_t & z_t \in A^c \end{cases} \quad (5.1)$$

$$(5.2)$$

con $\rho \in (-1, 1)$. Por otro, los modelos ARTIMA introducidos a lo largo de esta tesis, y que en su forma mas simple tienen la siguiente expresión,

$$(1 - L)y_t = \begin{cases} \varepsilon_t - \theta_1 \varepsilon_{t-1} & z_t \in A \\ \varepsilon_t - \theta_2 \varepsilon_{t-1} & z_t \in A^c \end{cases} \quad (5.3)$$

$$(5.4)$$

Respecto a los modelos ARTIMA, a lo largo de este capítulo, nos centraremos en el caso del tamaño de la perturbación como variable threshold, esto es, $z_t = \varepsilon_{t-1}$ y $A = \{|z_t| > r\}$, para $(\theta_1, \theta_2) = (0, 1)$ y $(1, 0)$. En el primer caso, las perturbaciones pequeñas serán transitorias y las grandes permanentes, y en el segundo, al revés. La extensión de los resultados a los otros casos estudiados en esta tesis es directa.

Larga memoria y agregación sectorial

En primer lugar estudiamos el caso de agregación sectorial. Siguiendo a Granger (1980), consideremos la agregación de N series AR(1),

$$x_{i,t} = \alpha_i x_{i,t-1} + \varepsilon_{i,t}$$

con $\varepsilon_{i,t}$ ruido blanco y varianza $var(\varepsilon_i)$ e independientes de α_i y $\varepsilon_{j,t}$ para todo i, j, t . Se puede probar que el espectro de la serie agregada $x_t = \sum_{i=1}^N x_{i,t}$ se puede aproximar por,

$$f_x(\varpi) \approx \frac{N}{2\pi} E(var(\varepsilon_{i,t})) \int \frac{1}{|1 - \alpha e^{i\varpi}|^2} \partial F(\alpha).$$

donde F es la función de distribución de α . Asumiendo que

$$\partial F(\alpha) = \frac{2}{B(p, q)} \alpha^{2p-1} (1 - \alpha^2)^{q-1} \partial \alpha \quad 0 \leq \alpha \leq 1,$$

con $p > 0$ y $q > 0$, la función de autocorrelación de x_t será

$$\gamma_x(k) = \frac{2}{B(p, q)} \int_0^1 \alpha^{2p+k-1} (1 - \alpha^2)^{q-1} \partial \alpha = A k^{1-q}.$$

De esta forma, siguiendo la primera definición de larga memoria, x_t será un proceso integrado de orden $1 - q/2$, $x_t \sim I(1 - q/2)$. A medida que q tiende a 0, esto es, más masa de probabilidad hay entorno a $\alpha = 1$, mayor es la persistencia de la serie. Como ya hemos comentado, Zaffaroni (1999) prueba que el resultado es el mismo si,

$$\partial F(\alpha) \sim c_q L((1 - \alpha)^q \quad \text{cuando} \quad \alpha \rightarrow 1^-.$$

Para extender esta argumentación a los modelos threshold, consideramos en primer lugar que las series $x_{i,t}$ son generadas por modelos TUR, con la siguiente expresión,

$$y_{i,t} = \begin{cases} y_{i,t-1} + \varepsilon_{i,t} & z_{i,t} \in A_i \\ \rho_i y_{i,t-1} + \varepsilon_{i,t} & z_{i,t} \in A_i^c, \end{cases}$$

o de forma más compacta,

$$y_{i,t} = \rho_{i,t} y_{i,t-1} + \varepsilon_{i,t},$$

con $\varepsilon_{i,t}$ ruido blanco independientes de ρ_i y A_i . Asumiendo que $z_{i,t}$ es *iid* y $E(\rho_{i,t} \varepsilon_{i,t-1} | \mathfrak{S}_{t-1}) = 0$, es fácil probar, usando la descomposición de Wold, que

$$y_{i,t} = \rho_{i,E} y_{i,t-1} + u_{i,t}$$

con $\rho_{i,E} = P(z_{i,t} \in A_i) + \rho_i P(z_{i,t} \in A_i^c)$ y $u_{i,t}$ ruido blanco. Para replicar los resultados del caso AR(1) lineal, hemos de justificar que la distribución de $\rho_{i,E}$, $f_{\rho_{i,E}}$, cumpla,:

$$f_{\rho_{i,E}(\alpha)} \sim c_q L((1 - \alpha)^q \quad \text{as} \quad \alpha \rightarrow 1^-$$

Dado que $z_{i,t}$ es *iid*, la distribución de $\rho_{i,E}$ dependerá solo de la distribución de $p_i = P(z_{i,t} \in A_i)$ y ρ_i . Dado que p_i es una probabilidad, la distribución Beta(p,q) se revela como la más adecuada. Asumiendo la independencia entre p_i y ρ_i tenemos que

$$f_{\rho_{i,E}}(\alpha) = \int_{-1}^1 f_p\left(\frac{\alpha - \rho}{1 - \rho}\right) f_\rho(\rho) \partial \rho \propto c_{p_E} (1 - \alpha)^q \quad \text{cuando} \quad \alpha \rightarrow 1^-$$

donde f_ρ es la densidad de ρ_i y $c_{pE} \propto \int_{-1}^1 (\frac{\alpha-\rho}{1-\rho})^p (1-\rho)^{-q} f_\rho(\rho) \partial\rho$, que suponemos existe. Por lo tanto, la agregación de procesos TUR también puede generar larga memoria, ofreciendo una justificación a las distribuciones Beta necesarias para el parámetro autoregresivo.

Para el caso en que las series $x_{i,t}$ sean generadas por procesos ARTIMA, como los descritos al inicio de la sección, se puede probar que a través de la descomposición de Wold,

$$(1 - L)x_{i,t} = u_{i,t} - \theta_i u_{i,t-1}$$

con $u_{i,t}$ ruido blanco. Por lo tanto, la representación lineal de los procesos ARTIMA, será un IMA. Como establece Granger, la agregación de procesos IMA es un IMA, y no es posible generar procesos que reproduzcan propiedades de larga memoria con $d < 1$. Por lo tanto, para estos procesos no será posible generar larga memoria a través de la agregación sectorial.

Larga memoria y no linealidades

Otros procesos que pueden originar larga memoria son los modelos de cambios estructurales infrecuentes. Consideremos el siguiente modelo, propuesto por Granger y Hyung (1999) y Diebold e Inoue (2001)

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t \\ \mu_t &= \mu_{t-1} + q_t v_t \\ q_t &= \begin{cases} 0 & \text{con probabilidad } p \\ 1 & \text{e.o.c} \end{cases} \end{aligned}$$

con ε_t , v_t y q_t iid. Si definimos $S_T = \sum_{t=1}^T y_t$, es fácil probar que $Var(S_T) = O(T^3 p)$. Por lo tanto si $p > 0$ y constante, $y_t \sim I(1)$. Sin embargo, si asumimos que $p = O(T^{2d-2})$, entonces $Var(S_T) = O(T^{2d+1})$, con lo que y_t reproducirá la propiedad de un proceso integrado $FI(d)$ en el orden de magnitud de las sumas parciales. Los valores de d de interés son

$d \in (-0.5, 1)$, por lo tanto para obtener alguna propiedad de larga memoria en procesos con cambios estructurales es necesario que la probabilidad de que haya un break tienda a 0 con el tamaño muestral, lo cual implica que en una serie de tamaño infinito el número de breaks debe ser finito.

En el caso de los modelos threshold, el parámetro que mide la probabilidad de cambio de régimen es el parámetro threshold. Así, para el caso TUR con z_t independiente de ε_t , sabemos que

$$\begin{aligned} Var(y_t) &= \sigma_\varepsilon^2 \frac{1 - \sigma_\rho^{2t}}{1 - \sigma_\rho^2} \simeq \frac{\sigma_\varepsilon^2}{1 - \sigma_\rho^2} \\ Cov(y_t, y_{t-s}) &= \sigma_\varepsilon^2 \rho_E^s \frac{1 - \sigma_\rho^{2t}}{1 - \sigma_\rho^{2t}} \simeq \rho_E^s \frac{\sigma_\varepsilon^2}{1 - \sigma_\rho^2} \\ \sigma_\rho^2 &= p + \rho^2(1 - p) \\ \rho_E &= p + \rho(1 - p) \end{aligned}$$

con $p = P(z_t \in A)$. Por lo tanto, es fácil comprobar que si p es fijo, $Var(S_T) = O(T)$ e y_t será un proceso $I(0)$. Por el contrario, supongamos que la probabilidad de cada régimen depende de T , $p(T)$. Dado que partimos de un proceso $I(0)$, para conseguir que el proceso genere propiedades de larga memoria será necesario que la probabilidad de estar en el régimen de raíz unitaria, $z_t \in A$, deberá tender a 1 con T , $p(T) = 1 - T^{-\delta}$ con $\delta > 0$. En este caso, $\sigma_\rho^2 = \sigma_\rho^2(T) = p(T) + \rho^2(1 - p(T)) \propto 1 - T^{-\delta}$ y $\rho_E = 1 - T^{-\delta}$, por lo tanto, $Var(S_T) \propto T \sigma_\varepsilon^2 T^{2\delta} = O(T^{1+2\delta})$ con lo que y_t tendría la propiedad de un proceso $I(d)$, con $d = \delta$, respecto del orden de convergencia de su suma parcial.

Para los procesos ARTIMA, descritos al principio de la sección, trabajaremos con $x_t = (1 - L)y_t$, de forma que si $x_t \sim I(d)$ entonces $y_t \sim I(d + 1)$. En concreto, es fácil probar

que,

$$\begin{aligned} Var(x_t) &= (1 + \theta_1^2)\sigma_\varepsilon^2 + (\theta_2^2 - \theta_1^2)\Lambda(r) \\ Cov(x_t, x_{t-s}) &= \begin{cases} -[\theta_1\sigma_\varepsilon^2 + (\theta_2 - \theta_1)\Lambda(r)] & si \quad s = 1 \\ 0 & si \quad s > 1 \end{cases} \end{aligned}$$

con $\Lambda(r) = E(\varepsilon_t^2 1(z_t \in A)) = E(\varepsilon_t^2 1(|\varepsilon_{t-1}| < r)) = \int_{-r}^r \varepsilon^2 f_\varepsilon(\varepsilon) d\varepsilon$. Por lo tanto, en el caso de $(\theta_1, \theta_2) = (0, 1)$, perturbaciones pequeñas transitorias, tenemos que

$$Var\left(\sum_{t=1}^T x_t\right) = T[\sigma_\varepsilon^2 - \Lambda(r)] + 2\Lambda(r)$$

Para el segundo caso, $(\theta_1, \theta_2) = (1, 0)$, perturbaciones grandes transitorias,

$$Var\left(\sum_{t=1}^T x_t\right) = T\Lambda(r) + 2[\sigma_\varepsilon^2 - \Lambda(r)]$$

Si r es constante, entonces en ambos casos $x_t \sim I(0)$, y por lo tanto $y_t \sim I(1)$. Para este caso, si queremos obtener alguna propiedad de larga memoria, la probabilidad de estar en el régimen transitorio, $\theta_i = 1$, debe converger a 1 con el tamaño muestral T . Así, para el primer caso, $(\theta_1, \theta_2) = (0, 1)$, supongamos que $r(T)$ es tal que $\sigma_\varepsilon^2 - \Lambda(r) = O(T^{-\delta})$. Bajo esta condición, $Var(\sum_{t=1}^T x_t) = O(T^{1-\delta})$ y por lo tanto $x_t \sim I(-\frac{\delta}{2})$ e $y_t \sim I(1 - \frac{\delta}{2})$. Por lo tanto, cuando las perturbaciones pequeñas son transitorias, r debe converger a ∞ , garantizando que la probabilidad de estar en el régimen transitorio converge a 1 cuando T tiende a ∞ . Para el segundo caso, $(\theta_1, \theta_2) = (1, 0)$, será necesario que $\Lambda(r) = O(T^{-\delta})$. Bajo esta condición, $Var(\sum_{t=1}^T x_t) = O(T^{1-\delta})$ y obtenemos los mismos resultados que en el caso anterior. Así, cuando las perturbaciones grandes son transitorias, r debe converger a 0, lo cuál garantiza que la probabilidad de estar en el régimen transitorio tiende a 1 cuando T tiende a ∞ .

Por lo tanto, para los procesos ARTIMA, como sucedía en el caso de breaks, la probabilidad de que haya una perturbación permanente debe converger a 0, esto es, cuando la muestra tiende a ∞ el número de perturbaciones permanentes debe ser finito. En otro

caso, no será posible reproducir ninguna propiedad de los procesos de larga memoria. Este resultado se invierte en el caso de los procesos *TUR*, ya que en el límite, $T \rightarrow \infty$, los procesos deberán ser integrados.

5.4. Estudio de Monte Carlo

En esta sección, trataremos de analizar la relación de los modelos ARTIMA-Shock con las propiedades de larga memoria en muestras finitas, a través de un estudio de Monte Carlo. Este análisis se centrará en tres aspectos fundamentales, que detallamos a continuación:

i) Comportamiento de la estimación del parámetro de larga memoria, d , en los procesos ARTIMA. Para ello usaremos el método de estimación semiparamétrico propuesto por Geweke y Porter-Hudak (1983), (GPH en adelante).

ii) Comportamiento de los contrastes de larga memoria en los procesos ARTIMA. Se analizarán dos estadísticos, el que se deriva de la estimación propuesta por GPH y el propuesto por Tanaka (1999). El análisis consistirá en el estudio del porcentaje de rechazo de las hipótesis nulas $d = 0$ y $d = 1$ de ambos contrastes. Cuando el porcentaje de rechazo de ambas hipótesis nulas sea elevado, indicará un elevado porcentaje de aceptación de larga memoria (entendida ésta como el caso $d \in (0, 1)$).

iii) Comportamiento del contraste de raíz unitaria propuesto por Dickey-Fuller (1979, 1981), (D-F en adelante).

Dado que el objetivo fundamental del análisis es comprobar la relación de la larga memoria con la no linealidad, y en particular, con los modelos ARTIMA-Shock-Size presentados en los capítulos anteriores, a lo largo del mismo utilizaremos dos modelos lineales

como marco de referencia, los modelos ARIMA y los modelos ARFIMA. Concretamente, se analizarán los tres comportamientos descritos anteriormente en los modelos ARIMA y ARFIMA que mejor se ajustan a los modelos ARTIMA-Shock en cada caso. Este marco de referencia permitirá discernir si este comportamiento se origina en el carácter no lineal del modelo, o por el contrario, es una característica compartida con los modelos lineales. Finalmente, este análisis se replicará para otro modelo no lineal, los modelos TUR, que como hemos visto son modelos threshold autoregresivos con un régimen de raíz unitaria. Concretamente los modelos analizados serán:

$$\textbf{ARTIMA-Shock-Size 1: } (1 - L)y_t = \varepsilon_t - \varepsilon_{t-1}1(|\varepsilon_{t-1}| \leq r). \quad (5.5)$$

$$\textbf{ARTIMA-Shock-Size 2: } (1 - L)y_t = \varepsilon_t - \varepsilon_{t-1}1(|\varepsilon_{t-1}| > r). \quad (5.6)$$

$$\textbf{TUR 1: } y_t = y_{t-1}1(|\varepsilon_{t-1}| > r) + 0,5y_{t-1}1(|\varepsilon_{t-1}| \leq r) + \varepsilon_t. \quad (5.7)$$

En el primer modelo, ARTIMA-Shock-Size 1, las perturbaciones pequeñas serán transitorias y las grandes permanentes. En el segundo modelo, ARTIMA-Shock-Size 2, el comportamiento es el opuesto. Y en el modelo TUR, todas las perturbaciones serán transitorias, como ya vimos en los capítulos anteriores. Esto marca la diferencia fundamental entre ambos tipos de modelos, ARTIMA y TUR. En principio, los modelos TUR serán procesos $I(0)$, mientras que los modelos ARTIMA serán $I(1)$. Así, la aproximación ARIMA a los modelos ARTIMA se hará a través del modelo:

$$(1 - L)y_t = u_t - \theta u_{t-1} \quad (5.8)$$

y para el modelo TUR, a través de

$$y_t = \phi y_{t-1} + u_t. \quad (5.9)$$

En la aproximación ARFIMA a los tres modelos, se utilizará la ecuación:

$$(1 - L)^d y_t = u_t. \quad (5.10)$$

Así, por ejemplo, para el caso ARTIMA-Shock-Size 1 generaremos a partir de $\{\varepsilon_t^i\}_{t=1}^T$ cada serie $\{y_t^i\}_{t=1}^T$ según la ecuación (5.5). A cada serie $\{y_t^i\}_{t=1}^T$ le ajustaremos los modelos (5.8) y (5.10), obteniendo $\hat{\theta}^i$ y \hat{d}^i . Con \hat{d}^i y θ , correspondiente a la representación de Wold de (5.5), y $\{\varepsilon_t^i\}_{t=1}^T$ generaremos las aproximaciones ARIMA y ARFIMA a través de las ecuaciones (5.8) y (5.10) respectivamente.

Cada estudio se realiza para B=1000 replicaciones y distintos tamaños muestrales, T=200, 500, 800, 1500. Finalmente, también se considerarán distintos valores del parámetro threshold, que regulará la probabilidad de estar en cada régimen, excepto en el caso del modelo ARTIMA-Shock-Size 1, perturbaciones grandes permanentes, en el que el parámetro a controlar será el número de perturbaciones permanentes. En todos los casos, la variables ε_t se obtienen de $N(0, 1)$ independientes.

En las Tablas 5.1, 5.2 y 5.3 se encuentran los resultados de la estimación GPH del parámetro d para los tres modelos. La Tabla 5.1 muestra los siguientes resultados para el modelo ARTIMA-Shock-Size 1. En la primera columna, el valor medio de r necesario para obtener el número de perturbaciones permanentes considerado. En la segunda columna, el número de perturbaciones permanentes, que serán, T, 4, 3, 2 y 1. En las siguientes dos columnas aparecen, la media del parámetro θ correspondiente a la representación de Wold de un proceso ARTIMA-Shock-Size 1, $E(\theta)$, y la media del valor realmente estimado, $\hat{\theta}$, si ajustamos un modelo ARIMA al proceso ARTIMA-Shock-Size 1. En las siguientes tres columnas, aparecen la media del valor estimado para el parámetro d en el proceso ARTIMA y en sus aproximaciones ARIMA y ARFIMA. Finalmente, en las últimas tres columnas aparece la media del valor estimado para el R^2 de la regresión necesaria para d en cada uno de los tres casos anteriormente descritos. En la Tabla 5.2 se obtienen los mismos resultados para el modelo ARTIMA-Shock-Size 2, aunque en este caso se fija r , presentando la media del porcentaje de perturbaciones permanentes. La Tabla 5.3, presenta los mismos resultados para el modelo TUR, y por lo tanto, el porcentaje de permanentes refleja el número de

veces que estamos en el régimen de raíz unitaria. Como cabía esperar, para el modelo ARTIMA-Shock-Size 1, el valor medio de r necesario para obtener un determinado número de perturbaciones permanentes aumenta con T . En el caso de modelo ARTIMA-Shock-Size 2, a medida que aumenta r , aumentará el porcentaje de perturbaciones permanentes. En el modelo TUR, a medida que r aumenta menor será el número de veces que estamos en el régimen de raíz unitaria.

Vistos los resultados de estas tres tablas, las conclusiones que se obtienen serían las siguientes. En primer lugar, que los valores estimados de d para el proceso ARTIMA-Shock-Size 1 se concentran en torno a $d = 1$, cosa que no sucede en los otros dos casos. También se puede observar que en todos los casos, el ajuste en la estimación de d es mejor cuanto más se acerca dicha estimación de d a 1. Finalmente, en cuanto a la comparación de los tres modelos con sus aproximaciones lineales, la propiedad de memoria representada a través del parámetro d es fielmente recogida por dichas aproximaciones.

En las Tablas 5.4, 5.5 y 5.6 se presentan los resultados del contraste de D-F Aumentado para los tres modelos objeto de estudio. En la tabla 5.4 aparecen los resultados para el modelo ARTIMA-Shock-Size 1 y los distintos números de perturbaciones permanentes y tamaños muestrales. La hipótesis nula será la existencia de raíz unitaria, $d = 1$, y se considerarán cuatro valores posibles para los retardos, $k = 0, 3, 6, 10$. La justificación en la utilización del contraste de D-F aumentado se debe a que la representación de Wold del proceso es un IMA, ecuación (5.8). Todos los contrastes se hacen al nivel de significación 0.05, y por tanto, dado que los procesos ARTIMA son $I(1)$, el porcentaje de rechazos de H_0 debería ser del 5 %. En las Tablas 5.5 y 5.6 se presentan los mismos resultados para los procesos ARTIMA-Shock-Size 2 y TUR respectivamente. Sin embargo, como en el análisis anterior, para estos dos casos se fija el valor del parámetro r . Además, para el proceso TUR, la hipótesis nula $d = 1$ es falsa, y por tanto, el porcentaje de rechazos debería converger al 100 %.

A la vista de los resultados las conclusiones son las siguientes. En primer lugar, la aproximación lineal ARFIMA refleja peor el comportamiento de los modelos no lineales que el modelo ARIMA, sobre todo para el caso de los modelos ARTIMA. Esto se debe a que, salvo para $d = 1$, el modelo ARFIMA no es un $I(1)$. También se puede observar que el modelo lineal ARIMA se comporta de forma muy similar a los tres modelos. Sin embargo, para el modelo ARTIMA de perturbaciones grandes permanentes, cuando sólo hay una perturbación permanente y el número de retardos introducidos en el estadístico de contraste es elevado, $k = 6, 10$, el comportamiento ante el contraste de DF Aumentado es sensiblemente superior al reproducido por el modelo lineal ARIMA. Por último, como cabía esperar, el contraste de D-F Aumentado funciona mejor para los modelos ARTIMA de perturbaciones grandes permanentes. Por ejemplo, para el tamaño $T = 500$ y $k = 10$ retardos, es suficiente que el 0.8 % de las perturbaciones más grandes sean permanentes para que el tamaño del contraste este cercano al 5 %. Para las perturbaciones pequeñas, este porcentaje asciende al 68 % de las perturbaciones. Este resultado refleja la importancia que en el contraste de D-F tiene la varianza del componente permanente. Para el modelo TUR, no hay justificación para el contraste de DF Aumentado, por lo tanto, en la comparación seleccionamos $k = 0$. Para el tamaño $T = 500$, será suficiente con que la probabilidad de estar en el régimen “transitorio” sea 0.08 para obtener una potencia del 80 %.

Por último, en las Tablas 5.7, 5.8 y 5.9 se presentan los resultados de los contrastes de GPH y Tanaka para las hipótesis nulas $d = 1$ y $d = 0$ y los tres modelos estudiados. Como se comentó anteriormente, cuando el porcentaje de rechazo de ambas hipótesis sea elevado, esto implicará un porcentaje elevado en la aceptación de larga memoria. El contraste de GPH utiliza la estimación del parámetro d presentada en las Tablas 5.1, 5.2 y 5.3 para construir el estadístico. Como el estimador, se tratará de un contraste semiparamétrico, esto es, aunque el modelo bajo la nula sea ARFIMA(p,d,q) no es necesaria la especificación y estimación de la parte ARMA(p,q). El estadístico será un t-ratio estándar construido a

partir de la ecuación de regresión usada en la estimación de d . Por el contrario, el contraste de Tanaka es un contraste paramétrico en el que es necesario la estimación de la parte ARMA(p,q). Así, si suponemos que la serie y_t sigue la ecuación (5.10) con

$$(1 - bL)u_t = (1 - aL)\varepsilon_t$$

y ε_t *iid*, el estadístico de Tanaka será:

$$S_T = T \sum_{k=1}^{T_1} \frac{1}{k} \hat{\rho}_k,$$

donde $\hat{\rho}_k$ es la autocorrelación de orden k de los residuos $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_{T-1}$, con

$$\hat{\varepsilon}_t = (1 - \hat{a}L)(1 - \hat{b}L)^{-1}(1 - L)^{d_0}y_t.$$

A la hora de calcular este estadístico para el modelo TUR, supondremos $a = 0$ y $b = \alpha \neq 0$. Sin embargo, para el caso de los modelos ARTIMA, supondremos $b = 0$ y $a = \alpha \neq 0$. En ambos casos, y bajo $H_0 : d = d_0$,

$$\frac{1}{\sqrt{T}}S_T \rightarrow N(0, w),$$

donde,

$$\hat{w}^2 = \frac{\pi^2}{6} - \frac{1 - \hat{\alpha}^2}{\hat{\alpha}^2} (\log(1 - \hat{\alpha}))^2.$$

Respecto al contraste de GPH y el modelo ARTIMA-Shock-Size 1, el porcentaje de rechazos de $d = 0$, esto es, la potencia del contraste, es muy cercana a 1 en todos los casos. Respecto al tamaño, porcentaje de rechazos de $d = 1$, el comportamiento es algo peor, empeorando al aumentar T y fundamentalmente para los casos de 1 ó 2 perturbaciones permanentes. En el peor de los casos, $T = 1500$ y 1 perturbación permanente, el 31 % de las veces aceptaremos la existencia de larga memoria basados en este contraste. En cuanto a la comparación de este comportamiento con el de su aproximación lineal ARIMA, conviene destacar que el comportamiento en ambos casos es muy similar, aunque el contraste en

este caso lineal parece mejorar levemente en potencia y empeorar en tamaño. Respecto a la aproximación ARFIMA, el porcentaje de rechazos de $H_0 : d = 1$ es superior al caso ARTIMA y ARIMA, aunque tiende a igualarse a medida que T aumenta y el número de perturbaciones permanentes disminuye.

Para el modelo ARTIMA de perturbaciones pequeñas permanentes, ARTIMA-Shock-Size 2, la potencia del contraste funciona igualmente bien, aunque el tamaño se comporta peor. En concreto, para el tamaño muestral $T = 800$ y un porcentaje de perturbaciones permanentes del 19 % aceptaremos la existencia de larga memoria en el 75 % de los casos. Como se vio en la sección anterior, los resultados reflejan que para aceptar la existencia de larga memoria, a medida que aumenta el tamaño muestral, debe disminuir el porcentaje de perturbaciones permanentes. Comparado con su aproximación ARIMA, el comportamiento del contraste es muy similar. Respecto a la aproximación ARFIMA los resultados son similares a los obtenidos para los modelos ARTIMA y ARIMA.

En el modelo TUR es donde el contraste de GPH se comporta peor. En este caso, la potencia del contraste se mide en el porcentaje de rechazos de la hipótesis $d = 1$, que es sensiblemente inferior a los casos anteriores. El tamaño del contraste vendrá dado por el porcentaje de rechazos de $d = 0$, que en general es muy superior al teórico del 5 %. En particular, para el tamaño $T = 800$ y una probabilidad de 0.86 de estar en el régimen de raíz unitaria aceptaremos la existencia de larga memoria en el 87 % de los casos. De nuevo, este comportamiento es reproducido de forma similar por la aproximación lineal ARFIMA, y fielmente por la aproximación ARIMA. No sucede así para el contraste de Tanaka. Este estadístico presenta un mejor tamaño, empeorando levemente en la potencia, siendo esta mejora en el tamaño y empeoramiento en la potencia mucho mayor en el caso de la aproximación lineal ARIMA que en el TUR.

Para el caso del modelo ARTIMA con perturbaciones pequeñas permanentes, el comportamiento del contraste de Tanaka mejora ostensiblemente, siendo la potencia muy cercana

a 1 y el tamaño a 0.05 en la mayoría de los casos. La mejoría es sensiblemente superior para la aproximación ARIMA. Sin embargo, esta mejora, fundamentalmente en el tamaño, se hace a costa de una disminución de la potencia cuando el porcentaje de permanentes es exactamente 0, esto es, $d = 0$ y sin embargo el porcentaje de rechazos de $d = 1$ no supera el 25 %. Estas conclusiones son extrapolables al caso de perturbaciones grandes permanentes.

Por lo tanto, el contraste de Tanaka parece el más adecuado para rechazar la existencia de larga memoria en el contexto de modelos no lineales, aunque conlleva un coste en términos de potencia cuando el proceso verdadero es un ruido blanco. Esta disminución de la potencia se debe a que a pesar de ser ruido blanco, ajustamos a la serie un ARIMA(0,1,1). Por lo tanto, como cabía esperar al tratarse de un contraste paramétrico, si la especificación es correcta, el contraste funciona mejor que el de GPH. Sin embargo, cuando la especificación es errónea, caso $r = 0$, el contraste presenta un comportamiento peor que el de GPH. Finalmente, todas las aproximaciones lineales se comportan de forma similar a sus modelos threshold, exceptuando el modelo TUR para el contraste de Tanaka.

Cuadro 5.1: Resultados de la Estimación
ARTIMA: Perturbaciones Grandes Permanentes

T=200		ARIMA		d			R^2		
\bar{r}	n°Perman	$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
0	T	0	0,003 (0,073)	0,988 (0,206)	0,988 (0,206)	0,930 (0,309)	0,661 (0,165)	0,661 (0,165)	0,646 (0,234)
2.322	4	0.666	0,679 (0,057)	0,903 (0,224)	0,889 (0,223)	0,881 (0,308)	0,609 (0,171)	0,600 (0,182)	0,595 (0,226)
2.425	3	0.699	0,713 (0,055)	0,876 (0,235)	0,867 (0,227)	0,858 (0,313)	0,600 (0,178)	0,585 (0,186)	0,581 (0,229)
2.568	2	0.739	0,754 (0,057)	0,830 (0,277)	0,832 (0,234)	0,814 (0,353)	0,571 (0,195)	0,559 (0,192)	0,555 (0,245)
2.814	1	0.799	0,807 (0,042)	0,771 (0,297)	0,752 (0,249)	0,763 (0,386)	0,510 (0,200)	0,500 (0,205)	0,526 (0,243)

T=500		ARIMA		d			R^2		
\bar{r}	n°Perman	$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
0	T	0	0,003 (0,073)	0,990 (0,187)	0,990 (0,187)	0,934 (0,276)	0,596 (0,154)	0,596 (0,154)	0,586 (0,221)
2.322	4	0.759	0,764 (0,031)	0,907 (0,207)	0,893 (0,201)	0,882 (0,275)	0,543 (0,167)	0,532 (0,169)	0,535 (0,215)
2.425	3	0.784	0,788 (0,030)	0,894 (0,198)	0,871 (0,205)	0,873 (0,270)	0,537 (0,166)	0,517 (0,172)	0,527 (0,213)
2.568	2	0.816	0,820 (0,031)	0,843 (0,249)	0,833 (0,211)	0,828 (0,313)	0,499 (0,174)	0,492 (0,174)	0,493 (0,224)
2.814	1	0.861	0,861 (0,022)	0,772 (0,250)	0,749 (0,224)	0,769 (0,337)	0,429 (0,170)	0,435 (0,181)	0,451 (0,221)

T=800		ARIMA		d			R^2		
\bar{r}	n°Perman	$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
0	T	0	0,001 (0,033)	0,990 (0,187)	0,990 (0,187)	0,934 (0,276)	0,596 (0,154)	0,596 (0,154)	0,586 (0,221)
2.806	4	0.799	0,800 (0,021)	0,907 (0,207)	0,893 (0,201)	0,882 (0,275)	0,543 (0,167)	0,532 (0,169)	0,535 (0,215)
2.897	3	0.810	0,822 (0,021)	0,894 (0,198)	0,871 (0,205)	0,873 (0,270)	0,537 (0,166)	0,517 (0,172)	0,527 (0,213)
3.025	2	0.846	0,849 (0,022)	0,843 (0,249)	0,833 (0,211)	0,828 (0,313)	0,499 (0,174)	0,492 (0,174)	0,493 (0,224)
3.246	1	0.884	0,884 (0,017)	0,772 (0,250)	0,749 (0,224)	0,769 (0,337)	0,429 (0,170)	0,435 (0,181)	0,451 (0,221)

T=1500		ARIMA		d			R^2		
\bar{r}	n°Perman	$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
0	T	0	0,001 (0,025)	1,004 (0,149)	1,004 (0,149)	0,963 (0,218)	0,593 (0,132)	0,593 (0,132)	0,590 (0,194)
3.005	4	0.843	0,843 (0,015)	0,924 (0,154)	0,925 (0,159)	0,922 (0,202)	0,543 (0,136)	0,540 (0,141)	0,542 (0,171)
3.093	3	0.860	0,860 (0,015)	0,900 (0,162)	0,907 (0,162)	0,902 (0,211)	0,523 (0,138)	0,527 (0,142)	0,525 (0,175)
3.212	2	0.882	0,882 (0,016)	0,864 (0,194)	0,874 (0,168)	0,872 (0,243)	0,496 (0,146)	0,504 (0,146)	0,504 (0,183)
3.425	1	0.912	0,912 (0,012)	0,801 (0,189)	0,799 (0,182)	0,819 (0,255)	0,433 (0,133)	0,452 (0,153)	0,458 (0,179)

Cuadro 5.2: Resultados de la Estimación
ARTIMA: Perturbaciones Pequeñas Permanentes

T=200	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
r=0.0	0.0	1.000	0.934	-0,005 (0,225)	-0,005 (0,225)	0,003 (0,420)	0,072 (0,090)	0,072 (0,090)	0,188 (0,186)
r=0.5	38.4	0.836	0.840	0,700 (0,243)	0,698 (0,248)	0,705 (0,337)	0,456 (0,202)	0,462 (0,204)	0,475 (0,247)
r=1.0	68.3	0.609	0.619	0,932 (0,217)	0,917 (0,217)	0,898 (0,285)	0,624 (0,170)	0,620 (0,175)	0,615 (0,220)
r=1.5	86.7	0.394	0.402	0,981 (0,215)	0,964 (0,209)	0,930 (0,281)	0,656 (0,159)	0,650 (0,167)	0,636 (0,223)
r=2.0	95.5	0.214	0.220	0,987 (0,209)	0,980 (0,207)	0,932 (0,293)	0,659 (0,166)	0,657 (0,166)	0,643 (0,227)

T=500	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
r=0.00	0.0	1.000	0.959	-0,012 (0,218)	-0,012 (0,218)	-0,015 (0,414)	0,052 (0,071)	0,052 (0,071)	0,147 (0,159)
r=0.25	19.8	0.937	0.930	0,471 (0,235)	0,454 (0,239)	0,466 (0,357)	0,239 (0,165)	0,235 (0,169)	0,260 (0,206)
r=0.50	38.4	0.836	0.838	0,822 (0,220)	0,806 (0,213)	0,806 (0,299)	0,477 (0,176)	0,474 (0,175)	0,473 (0,220)
r=1.00	68.3	0.610	0.614	0,965 (0,199)	0,956 (0,189)	0,926 (0,246)	0,574 (0,158)	0,574 (0,160)	0,574 (0,209)
r=1.50	86.7	0.395	0.398	0,994 (0,193)	0,980 (0,186)	0,946 (0,255)	0,596 (0,152)	0,590 (0,155)	0,586 (0,211)

T=800	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
r=0.00	0.0	1.000	0.969	-0,011 (0,183)	-0,011 (0,183)	-0,017 (0,351)	0,040 (0,053)	0,040 (0,053)	0,119 (0,130)
r=0.10	8.0	0.983	0.965	0,142 (0,192)	0,138 (0,189)	0,132 (0,327)	0,068 (0,079)	0,064 (0,075)	0,123 (0,131)
r=0.25	19.8	0.938	0.934	0,577 (0,202)	0,568 (0,211)	0,572 (0,313)	0,324 (0,168)	0,319 (0,168)	0,322 (0,205)
r=0.50	38.4	0.836	0.838	0,888 (0,173)	0,866 (0,184)	0,868 (0,232)	0,541 (0,148)	0,531 (0,155)	0,536 (0,192)
r=0.75	54.7	0.724	0.727	0,959 (0,169)	0,940 (0,173)	0,927 (0,217)	0,587 (0,142)	0,582 (0,146)	0,582 (0,186)

T=1500	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	ARTIMA	ARIMA	ARFIMA	ARTIMA	ARIMA	ARFIMA
r=0.00	0.0	1.000	0.977	0,007 (0,159)	0,007 (0,159)	0,017 (0,309)	0,028 (0,038)	0,028 (0,038)	0,091 (0,105)
r=0.10	8.0	0.983	0.972	0,210 (0,176)	0,213 (0,181)	0,219 (0,289)	0,076 (0,079)	0,080 (0,082)	0,116 (0,119)
r=0.25	19.8	0.938	0.936	0,689 (0,191)	0,688 (0,190)	0,705 (0,270)	0,375 (0,156)	0,375 (0,155)	0,377 (0,186)
r=0.50	38.3	0.836	0.837	0,927 (0,157)	0,931 (0,157)	0,923 (0,197)	0,543 (0,134)	0,544 (0,140)	0,543 (0,172)
r=0.75	54.7	0.724	0.726	0,974 (0,153)	0,979 (0,151)	0,957 (0,191)	0,574 (0,138)	0,578 (0,134)	0,577 (0,174)

Cuadro 5.3: Resultados de la Estimación

TUR

T=200	% Perman	ARIMA		d			R^2		
		$E(\phi)$	$\hat{\phi}$	TUR	ARIMA	ARFIMA	TUR	ARIMA	ARFIMA
r=0.01	99.2	0.996	0.970	0,990 (0,218)	0,998 (0,213)	0,919 (0,306)	0,647 (0,160)	0,650 (0,155)	0,638 (0,238)
r=0.10	92.0	0.960	0.938	0,810 (0,239)	0,822 (0,222)	0,783 (0,370)	0,517 (0,182)	0,535 (0,177)	0,528 (0,254)
r=0.17	86.5	0.932	0.915	0,676 (0,243)	0,687 (0,225)	0,668 (0,408)	0,430 (0,190)	0,446 (0,188)	0,450 (0,259)
r=0.25	80.2	0.901	0.891	0,547 (0,234)	0,562 (0,225)	0,546 (0,414)	0,348 (0,186)	0,360 (0,189)	0,374 (0,250)
r=0.50	61.7	0.808	0.829	0,314 (0,237)	0,318 (0,227)	0,314 (0,421)	0,188 (0,158)	0,191 (0,156)	0,263 (0,221)

T=500	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	TUR	ARIMA	ARFIMA	TUR	ARIMA	ARFIMA
r=0.01	99.2	0.996	0.986	0,989 (0,207)	0,985 (0,208)	0,921 (0,302)	0,556 (0,143)	0,550 (0,149)	0,570 (0,238)
r=0.10	92.1	0.960	0.952	0,751 (0,225)	0,770 (0,15)	0,730 (0,360)	0,415 (0,166)	0,421 (0,160)	0,419 (0,230)
r=0.17	86.5	0.933	0.929	0,586 (0,210)	0,599 (0,217)	0,572 (0,370)	0,309 (0,159)	0,315 (0,158)	0,316 (0,216)
r=0.25	80.3	0.901	0.904	0,439 (0,216)	0,452 (0,213)	0,425 (0,381)	0,217 (0,144)	0,222 (0,144)	0,243 (0,197)
r=0.50	61.7	0.806	0.841	0,193 (0,210)	0,203 (0,212)	0,181 (0,392)	0,087 (0,090)	0,092 (0,096)	0,160 (0,158)

T=800	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	TUR	ARIMA	ARFIMA	TUR	ARIMA	ARFIMA
r=0.01	99.2	0.996	0.990	0,987 (0,170)	0,986 (0,175)	0,934 (0,250)	0,578 (0,128)	0,572 (0,130)	0,589 (0,210)
r=0.10	92.0	0.960	0.955	0,669 (0,189)	0,678 (0,182)	0,672 (0,324)	0,378 (0,153)	0,384 (0,144)	0,387 (0,214)
r=0.17	86.5	0.932	0.931	0,481 (0,185)	0,488 (0,184)	0,482 (0,337)	0,252 (0,140)	0,254 (0,138)	0,272 (0,202)
r=0.25	80.2	0.901	0.907	0,338 (0,186)	0,343 (0,183)	0,336 (0,342)	0,159 (0,118)	0,160 (0,118)	0,201 (0,178)
r=0.50	61.6	0.808	0.843	0,142 (0,176)	0,134 (0,184)	0,135 (0,337)	0,063 (0,068)	0,062 (0,070)	0,131 (0,138)

T=1500	% Perman	ARIMA		d			R^2		
		$E(\theta)$	$\hat{\theta}$	TUR	ARIMA	ARFIMA	TUR	ARIMA	ARFIMA
r=0.01	99.2	0.996	0.992	0,988 (0,160)	0,984 (0,161)	0,945 (0,216)	0,533 (0,112)	0,530 (0,116)	0,572 (0,196)
r=0.10	92.2	0.960	0.958	0,595 (0,171)	0,606 (0,163)	0,603 (0,289)	0,298 (0,128)	0,303 (0,124)	0,311 (0,185)
r=0.17	86.5	0.932	0.934	0,389 (0,167)	0,403 (0,161)	0,398 (0,290)	0,165 (0,104)	0,173 (0,106)	0,193 (0,159)
r=0.25	80.3	0.901	0.910	0,252 (0,163)	0,266 (0,159)	0,260 (0,290)	0,092 (0,081)	0,097 (0,082)	0,134 (0,132)
r=0.50	61.7	0.809	0.845	0,093 (0,160)	0,098 (0,157)	0,102 (0,291)	0,038 (0,046)	0,037 (0,045)	0,091 (0,101)

Cuadro 5.4: Resultados del Contraste de DF Aumentado

ARTIMA: Perturbaciones Grandes Permanentes

% Rechazos $H_0 : d = 1$												
T=200	ARTIMA				ARIMA				ARFIMA			
n°Perma	0	3	6	10	0	3	6	10	0	3	6	10
T	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.249	0.164	0.124	0.106
4	0.922	0.155	0.064	0.060	0.903	0.183	0.073	0.057	0.353	0.207	0.164	0.111
3	0.956	0.206	0.081	0.079	0.940	0.255	0.087	0.061	0.378	0.222	0.167	0.137
2	0.998	0.280	0.153	0.138	0.964	0.346	0.125	0.073	0.425	0.242	0.200	0.162
1	1.000	0.321	0.174	0.169	0.996	0.541	0.228	0.107	0.476	0.315	0.244	0.205

% Rechazos $H_0 : d = 1$												
T=500	ARTIMA				ARIMA				ARFIMA			
n°Perma	0	3	6	10	0	3	6	10	0	3	6	10
T	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.259	0.177	0.149	0.124
4	0.997	0.386	0.120	0.065	0.986	0.415	0.160	0.066	0.360	0.263	0.222	0.182
3	1.000	0.464	0.131	0.073	0.995	0.490	0.198	0.080	0.391	0.288	0.238	0.191
2	1.000	0.616	0.212	0.127	0.998	0.633	0.274	0.111	0.460	0.344	0.291	0.233
1	1.000	0.883	0.269	0.160	1.000	0.825	0.471	0.216	0.578	0.429	0.356	0.294

% Rechazos $H_0 : d = 1$												
T=800	ARTIMA				ARIMA				ARFIMA			
n°Perma	0	3	6	10	0	3	6	10	0	3	6	10
T	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.272	0.212	0.177	0.150
4	1.000	0.554	0.232	0.089	0.999	0.600	0.236	0.092	0.351	0.264	0.227	0.176
3	1.000	0.657	0.277	0.108	1.000	0.634	0.278	0.105	0.402	0.297	0.242	0.202
2	1.000	0.742	0.347	0.140	1.000	0.781	0.437	0.170	0.445	0.335	0.282	0.228
1	1.000	0.998	0.438	0.182	1.000	0.926	0.638	0.338	0.608	0.471	0.399	0.333

% Rechazos $H_0 : d = 1$												
T=1500	ARTIMA				ARIMA				ARFIMA			
n°Perma	0	3	6	10	0	3	6	10	0	3	6	10
T	0.049	0.049	0.049	0.049	0.050	0.049	0.050	0.050	0.264	0.192	0.164	0.143
4	1.000	0.734	0.363	0.143	1.000	0.787	0.420	0.181	0.398	0.297	0.261	0.232
3	1.000	0.835	0.459	0.157	1.000	0.858	0.505	0.239	0.470	0.354	0.305	0.283
2	1.000	0.935	0.598	0.245	1.000	0.923	0.634	0.346	0.504	0.399	0.347	0.326
1	1.000	1.000	0.879	0.295	1.000	0.980	0.838	0.551	0.666	0.544	0.485	0.420

Cuadro 5.5: Resultados del Contraste de DF Aumentado

ARTIMA: Perturbaciones Pequeñas Permanentes

	% Rechazos $H_0 : d = 1$											
	ARTIMA				ARIMA				ARFIMA			
T=200	0	3	6	10	0	3	6	10	0	3	6	10
r=0.0	1.000	1.000	1.000	0.995	1.000	1.000	1.000	0.995	1.000	0.991	0.952	0.864
r=0.5	1.000	0.669	0.296	0.127	1.000	0.680	0.294	0.129	0.669	0.452	0.327	0.248
r=1.0	0.837	0.117	0.050	0.056	0.826	0.134	0.055	0.056	0.329	0.184	0.135	0.095
r=1.5	0.401	0.055	0.043	0.055	0.408	0.058	0.046	0.050	0.271	0.150	0.129	0.102
r=2.0	0.175	0.048	0.053	0.061	0.177	0.050	0.051	0.050	0.245	0.160	0.126	0.115

	% Rechazos $H_0 : d = 1$											
	ARTIMA				ARIMA				ARFIMA			
T=500	0	3	6	10	0	3	6	10	0	3	6	10
r=0.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.989
r=0.25	1.000	0.998	0.951	0.711	1.000	1.000	0.947	0.717	0.942	0.855	0.807	0.714
r=0.50	1.000	0.697	0.320	0.126	1.000	0.724	0.338	0.130	0.521	0.397	0.342	0.262
r=1.00	0.832	0.137	0.069	0.045	0.818	0.131	0.068	0.052	0.295	0.198	0.164	0.136
r=1.50	0.397	0.062	0.054	0.054	0.378	0.058	0.053	0.049	0.264	0.190	0.156	0.124

	% Rechazos $H_0 : d = 1$											
	ARTIMA				ARIMA				ARFIMA			
T=800	0	3	6	10	0	3	6	10	0	3	6	10
r=0.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
r=0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998	0.996
r=0.25	1.000	1.000	0.958	0.761	1.000	1.000	0.960	0.758	0.909	0.852	0.793	0.727
r=0.50	1.000	0.778	0.369	0.146	1.000	0.745	0.372	0.141	0.467	0.332	0.282	0.229
r=0.75	0.981	0.342	0.128	0.073	0.980	0.334	0.104	0.062	0.313	0.218	0.182	0.141

	% Rechazos $H_0 : d = 1$											
	ARTIMA				ARIMA				ARFIMA			
T=1500	0	3	6	10	0	3	6	10	0	3	6	10
r=0.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
r=0.10	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.998
r=0.25	1.000	1.000	0.956	0.771	1.000	0.999	0.921	0.618	0.837	0.753	0.699	0.655
r=0.50	1.000	0.752	0.363	0.163	1.000	0.757	0.395	0.165	0.405	0.317	0.269	0.233
r=0.75	0.987	0.327	0.108	0.062	0.975	0.345	0.111	0.066	0.306	0.215	0.178	0.160

TUR

[illegible]

Cuadro 5.7: Resultados de los Contrastes de GPH y Tanaka

ARTIMA: Perturbaciones Grandes Permanentes

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
T=200	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
n°Perma	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T	0.988	1.000	0.988	1.000	0.945	0.991	0.090	0.034	0.090	0.034	0.186	0.163
4	0.979	0.998	0.982	1.000	0.937	0.986	0.172	0.050	0.175	0.054	0.240	0.206
3	0.976	0.991	0.976	0.999	0.937	0.980	0.195	0.063	0.203	0.072	0.249	0.244
2	0.943	0.958	0.969	0.998	0.897	0.949	0.230	0.078	0.257	0.091	0.288	0.233
1	0.902	0.924	0.934	0.976	0.868	0.918	0.242	0.114	0.383	0.111	0.336	0.299

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
T=500	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
n°Perma	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T	0.996	1.000	0.996	1.000	0.966	0.999	0.063	0.033	0.063	0.033	0.180	0.278
4	0.988	0.998	0.989	1.000	0.959	0.997	0.147	0.059	0.168	0.068	0.212	0.400
3	0.987	0.999	0.986	1.000	0.959	0.995	0.150	0.062	0.191	0.070	0.221	0.428
2	0.961	0.998	0.982	1.000	0.920	0.976	0.221	0.062	0.239	0.070	0.269	0.451
1	0.931	0.981	0.957	1.000	0.893	0.963	0.263	0.068	0.393	0.077	0.344	0.598

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
T=800	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
n°Perma	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T	0.996	1.000	0.996	1.000	0.966	0.999	0.063	0.033	0.063	0.033	0.180	0.278
4	0.988	0.998	0.989	1.000	0.959	0.997	0.147	0.059	0.168	0.068	0.212	0.400
3	0.987	0.999	0.986	1.000	0.959	0.995	0.150	0.062	0.191	0.070	0.221	0.428
2	0.961	0.998	0.982	1.000	0.920	0.976	0.221	0.062	0.239	0.070	0.269	0.451
1	0.931	0.981	0.957	1.000	0.893	0.963	0.263	0.068	0.393	0.077	0.344	0.598

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
T=1500	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
n°Perma	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T	1.000	1.000	1.000	1.000	0.997	1.000	0.069	0.037	0.069	0.037	0.165	0.349
4	0.999	1.000	1.000	1.000	0.998	1.000	0.143	0.059	0.167	0.059	0.194	0.518
3	0.999	1.000	1.000	1.000	0.997	1.000	0.183	0.050	0.193	0.054	0.235	0.569
2	0.982	0.998	1.000	1.000	0.978	0.991	0.225	0.057	0.260	0.054	0.268	0.606
1	0.978	0.991	0.999	1.000	0.988	0.963	0.310	0.066	0.398	0.058	0.348	0.767

Cuadro 5.8: Resultados de los Contrastes de GPH y Tanaka

ARTIMA: Perturbaciones Pequeñas Permanentes

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
T=200	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
r=0.0	0.057	0.007	0.057	0.007	0.214	0.203	0.994	0.252	0.994	0.252	0.930	0.385
r=0.5	0.904	0.954	0.901	0.957	0.854	0.952	0.445	0.131	0.451	0.124	0.446	0.365
r=1.0	0.986	1.000	0.988	1.000	0.955	0.994	0.141	0.058	0.144	0.048	0.205	0.204
r=1.5	0.994	1.000	0.986	1.000	0.963	0.997	0.101	0.051	0.103	0.047	0.173	0.167
r=2.0	0.988	1.000	0.986	1.000	0.950	0.996	0.084	0.052	0.094	0.057	0.176	0.157

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
T=500	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
r=0.00	0.062	0.006	0.062	0.006	0.202	0.255	1.000	0.159	1.000	0.159	0.945	0.400
r=0.25	0.722	0.964	0.711	0.957	0.661	0.911	0.804	0.101	0.831	0.101	0.706	0.474
r=0.50	0.979	1.000	0.974	1.000	0.932	0.996	0.267	0.076	0.293	0.071	0.322	0.527
r=1.00	0.993	1.000	0.996	1.000	0.979	1.000	0.092	0.065	0.097	0.063	0.161	0.324
r=1.50	0.997	1.000	0.997	1.000	0.981	1.000	0.081	0.067	0.071	0.062	0.171	0.286

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
T=800	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
r=0.00	0.058	0.007	0.058	0.007	0.202	0.268	1.000	0.152	1.000	0.152	0.980	0.408
r=0.10	0.257	0.496	0.235	0.533	0.335	0.548	0.997	0.142	0.999	0.135	0.968	0.208
r=0.25	0.928	0.993	0.902	1.000	0.813	0.987	0.751	0.103	0.773	0.102	0.687	0.687
r=0.50	0.995	1.000	0.994	1.000	0.985	0.999	0.195	0.072	0.249	0.075	0.272	0.557
r=0.75	0.999	1.000	0.998	1.000	0.993	1.000	0.106	0.059	0.130	0.058	0.185	0.383

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	ARTIMA		ARIMA		ARFIMA		ARTIMA		ARIMA		ARFIMA	
T=1500	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
r=0.00	0.049	0.015	0.049	0.015	0.235	0.338	1.000	0.128	1.000	0.128	0.989	0.386
r=0.10	0.405	0.884	0.431	0.887	0.469	0.758	0.997	0.088	0.995	0.094	0.962	0.185
r=0.25	0.980	1.000	0.989	1.000	0.934	0.999	0.632	0.063	0.627	0.074	0.533	0.815
r=0.50	0.999	1.000	1.000	1.000	0.997	1.000	0.149	0.054	0.156	0.062	0.199	0.520
r=0.75	1.000	1.000	1.000	1.000	0.997	1.000	0.082	0.067	0.088	0.064	0.147	0.411

Cuadro 5.9: Resultados de los Contrastes de GPH y Tanaka

TUR

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	TUR		ARIMA		ARFIMA		TUR		ARIMA		ARFIMA	
	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T=200												
r=0.01	0.992	0.071	0.995	0.067	0.951	0.321	0.094	0.054	0.078	0.055	0.203	0.209
r=0.10	0.957	0.120	0.962	0.062	0.878	0.120	0.285	0.124	0.258	0.099	0.347	0.384
r=0.17	0.900	0.150	0.922	0.047	0.792	0.061	0.468	0.203	0.456	0.133	0.478	0.564
r=0.25	0.812	0.182	0.824	0.034	0.707	0.050	0.683	0.317	0.662	0.169	0.589	0.769
r=0.50	0.483	0.258	0.489	0.019	0.501	0.139	0.914	0.623	0.917	0.350	0.786	0.951

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	TUR		ARIMA		ARFIMA		TUR		ARIMA		ARFIMA	
	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T=500												
r=0.01	0.997	0.071	0.994	0.065	0.954	0.408	0.089	0.061	0.091	0.053	0.199	0.270
r=0.10	0.954	0.142	0.961	0.049	0.867	0.163	0.368	0.275	0.328	0.165	0.400	0.663
r=0.17	0.880	0.201	0.873	0.035	0.750	0.209	0.661	0.498	0.628	0.286	0.576	0.888
r=0.25	0.719	0.303	0.740	0.024	0.611	0.354	0.851	0.704	0.856	0.453	0.737	0.971
r=0.50	0.287	0.496	0.315	0.015	0.381	0.476	0.986	0.976	0.982	0.961	0.898	0.999

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	TUR		ARIMA		ARFIMA		TUR		ARIMA		ARFIMA	
	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T=800												
r=0.01	0.999	0.069	0.999	0.076	0.982	0.378	0.079	0.069	0.082	0.064	0.194	0.295
r=0.10	0.971	0.146	0.973	0.047	0.883	0.180	0.619	0.415	0.604	0.281	0.547	0.858
r=0.17	0.874	0.252	0.862	0.042	0.733	0.467	0.896	0.714	0.897	0.521	0.759	0.980
r=0.25	0.645	0.408	0.654	0.033	0.572	0.610	0.975	0.903	0.973	0.801	0.874	0.998
r=0.50	0.253	0.669	0.240	0.018	0.360	0.534	1.000	0.998	0.997	0.999	0.963	1.000

	% Rechazos $H_0 : d = 0$						% Rechazos $H_0 : d = 1$					
	TUR		ARIMA		ARFIMA		TUR		ARIMA		ARFIMA	
	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA	GPH	TA
T=1500												
r=0.01	1.000	0.061	1.000	0.059	0.992	0.388	0.74	0.101	0.081	0.073	0.187	0.332
r=0.10	0.969	0.185	0.969	0.033	0.884	0.304	0.806	0.682	0.797	0.505	0.663	0.973
r=0.17	0.785	0.378	0.825	0.030	0.681	0.687	0.982	0.948	0.980	0.873	0.884	0.998
r=0.25	0.508	0.593	0.533	0.023	0.526	0.772	0.995	0.997	0.997	0.995	0.955	1.000
r=0.50	0.152	0.798	0.169	0.013	0.310	0.532	1.000	1.000	1.000	1.000	0.987	1.000

5.5. Conclusiones

En este capítulo se lleva a cabo un estudio detallado de la relación existente entre los modelos asimétricos en persistencia, ARTIMA y TUR, con los procesos de larga memoria ARFIMA.

En una primera etapa se estudian las relaciones teóricas entre estos modelos asimétricos y determinadas propiedades de los modelos de larga memoria. En concreto, se obtiene que la agregación sectorial de modelos TUR puede generar procesos ARFIMA, siendo esto imposible para el caso ARTIMA. Sin embargo, cuando el parámetro threshold de ambos modelos no lineales depende del tamaño muestral, la varianza de la suma parcial de dichos procesos puede presentar propiedades de larga memoria.

Finalmente, para comprobar estos resultados, se realiza un estudio de Monte Carlo, en el que además se compara el comportamiento, ante determinados contrastes de larga memoria, de estos procesos asimétricos con sus aproximaciones lineales ARIMA Y ARFIMA. En general podemos concluir que el contraste de Tanaka (1999) es más robusto a este tipo de asimetrías, aunque su resultado depende crucialmente de la correcta especificación del modelo. Además, exceptuando para este contraste y el modelo TUR, ambas aproximaciones lineales presentan comportamientos similares a los procesos asimétricos.

5.6. Bibliografía

Beran, J. (1994): *Statistics for Long-Memory Processes*. Chapman and Hall, New York.

Cioczek-Georges, R. y Mandelbrot, B. (1995): “A class of micropulses and antipersistent fractional Brownian motion”. *Stochastic Processes and Their Applications*, Vol. **60**, pag. 1-18.

Chambers, M.J. (1998): “Long memory and aggregation in macroeconomic time series”. *International Economic Review*, Vol. **39**, pag. 1053-1072.

Davidson, J. y Sibbertsen, P. (2004): “Generating schemes for long memory processes: regimes, aggregation and linearity”. Forthcoming in *Journal of Econometrics*.

Diebold, F. e Inoue, A. (2001): “Long memory and regime switching”. *Journal of Econometrics*, Vol. **105**, pag. 131-159.

Engle, R.F. y Smith, A.D. (1999): “Stochastic Permanent Breaks”. *The Review of Economics and Statistics*, Vol. **81**, 553-574.

Geweke, J. y Porter-Hudak, S. (1983): “The estimation and application of long memory time series models”. *Journal of Time Series Analysis*, Vol. **4**, pag. 221-238.

González, M. y Gonzalo, J. (1998): “Threshold Unit Root Processes”. Working Paper.

Gourieroux, C. y Jasiak, J. (2001): “Memory and infrequent breaks”. *Economics Letters*, Vol. **70**, pag. 29-41.

Granger, C.W.J. (1980): “Long memory relationships and the aggregation of dynamic models”. *Journal of Econometrics*, Vol. **14**, pag. 228-238.

Granger, C.W.J. y Hyung, N. (1999): “Occasional structural breaks and long memory with an application to the *S&P* 500 absolute stock returns”. *Journal of Empirical Finance*, Vol. **11**, pag. 399-421.

Granger, C.W.J. y Teräsvirta, T. (1999): “Simple nonlinear time series model with misleading linear properties”. *Economics Letters*, Vol. **62**, pag. 161-165.

Lippi, M. y Zaffaroni, P. (1999): “Contemporaneous Aggregation of linear dynamic models in large economies”. *Journal of Econometrics*, Vol. **120**, pag. 75-102.

Parke, W.R. (1999): “What is fractional integration?”. *The Review of Economics and Statistics*, Vol. **81**, pag. 632-638.

Taqqu, M., Willinger, W., y Sherman, R. (1997): “Proof of the fundamental result in self similar traffic modelling”. *Computer Communication Review*, Vol. **27**, pag. 5-23.

Tanaka, K. (1999): “The nonstationary fractional unit root”. *Econometric Theory*, Vol. **15**, pag. 549-582.

Capítulo 6

Consideraciones finales y líneas de investigación futura.

A la hora de realizar un análisis económico, uno de los aspectos más importantes es la identificación de las perturbaciones que conforman las distintas series económicas objeto del análisis. Esta Tesis introduce una nueva clase de modelos que permiten la identificación de las perturbaciones basada en sus propiedades de largo plazo (persistencia) y en criterios económicos más amplios, que a su vez sean contrastables. Esto facilita una identificación más endógena y flexible de las perturbaciones. Estos nuevos criterios surgirán de las propiedades de asimetría que presentan habitualmente las series económicas, y más concretamente, de las asimetrías en persistencia.

Las aportaciones al estudio de las asimetrías en las series económicas han sido muy amplias, y se remontan a los trabajos de Keynes (1936) o Burns y Mitchell (1946), centrados en el Business Cycle. Las características de las asimetrías en las series económicas, asociadas normalmente a la dinámica de la serie, han hecho que las principales aportaciones se hayan realizado en el contexto de los modelos no lineales. Sin embargo, el desarrollo teórico respecto a los modelos no lineales no ha sido todo lo amplio que cabría desear de-

bido, fundamentalmente, a dos causas: la complejidad de las relaciones entre las variables económicas, que dificulta la obtención de modelos concretos, y la complejidad de los propios modelos no lineales, que implica una mayor dificultad en la obtención de resultados de teoría asintótica. Este ha sido uno de los retos principales de la Tesis, de manera que uno de los objetivos centrales de la misma ha sido el de proponer modelos simples e intuitivos de los que obtener resultados de validez asintótica para los métodos de estimación y contraste utilizados en dichos modelos.

Las aportaciones concretas de la Tesis se han centrado tanto en el campo teórico como en el aplicado. Desde el punto de vista teórico, se presenta un nuevo método de identificación de las perturbaciones basado en un nuevo tipo de modelos threshold. Estos modelos, bajo determinadas condiciones, generarán asimetrías en la persistencia de las perturbaciones en términos de transitorio y permanente, dependiendo de algún criterio económico recogido a través de la variable threshold. Para estos modelos, se propone un método de estimación y de contraste de las hipótesis fundamentales, para los que se obtienen resultados de validez asintótica. Desde la perspectiva empírica, estos métodos se aplican a diferentes series económicas, como la tasa de desempleo, el producto nacional bruto y diferentes índices de precios. Finalmente, se analiza la relación existente entre los modelos no lineales presentados en la Tesis y las propiedades de larga memoria que presentan los modelos ARFIMA. En concreto, se encuentra que bajo determinadas condiciones estos modelos son capaces de generar algunas propiedades de los modelos ARFIMA. La inclusión de este análisis parecía particularmente importante por dos motivos: primero, porque los modelos ARTIMA suponen un enfoque alternativo al problema de rechazo de las dos hipótesis tradicionales respecto a la memoria de las series, $I(0)$ e $I(1)$, y segundo, por la ausencia de modelos económicos que produzcan la integración fraccional, tan habitual por otro lado, en las series económicas.

A continuación se resume brevemente las aportaciones que se han echo en cada uno de

los campos considerados. También se perfilan los campos futuros por los cuales se espera extender la investigación que se ha emprendido con este trabajo.

Identificación de las perturbaciones

Los métodos de identificación de perturbaciones desarrollados hasta el momento se han centrado fundamentalmente en el marco lineal. Estos métodos requieren de información a priori no contrastable. En los Capítulos 2 y 3 de esta Tesis se presentan una nueva clase de modelos no lineales de tipo threshold, que denominamos modelos ARTIMA, y que serán la base para la identificación de las perturbaciones. Esta se basará en el comportamiento a largo plazo de las perturbaciones y la asimetría de las mismas en dicho comportamiento. Sus principales características y ventajas se resumen a continuación:

- Los procesos ARTIMA son procesos autoregresivos integrados con un comportamiento threshold en la media móvil. Los modelos threshold presentan la ventaja de permitir una aproximación lineal por tramos a cualquier función no lineal suave. Una característica fundamental de los modelos ARTIMA es la presencia de una raíz unitaria en algún régimen de la media móvil. Como se prueba en los Capítulos 2 y 3, bajo determinadas condiciones, esto permitirá que las perturbaciones en dicho régimen sean transitorias, mientras que en el resto, son permanentes. Esto genera un nuevo marco en las asimetrías en persistencia, no estudiado hasta ahora, y que extiende estas asimetrías en la persistencia al caso de transitorio y permanente.
- Como variable threshold consideraremos las propiedades o características de alguna variable económica relevante, entre las que se incluye la propia perturbación, permitiendo una interpretación económica e intuitiva del modelo. Las propiedades en las que nos hemos centrado a lo largo de la tesis son el tamaño y el signo.
- La existencia de perturbaciones transitorias y permanentes en este contexto, elimina

la implicación negativa que suponen los modelos lineales de que todas las perturbaciones son, o transitorias o permanentes. Además, en el caso del paseo aleatorio, se elimina la imposición de que siempre haya una perturbación permanente.

Basado en los modelos ARTIMA y sus propiedades, en el Capítulo 4 se presenta un nuevo método de descomposición de las series en su componente permanente y transitorio. Las principales aportaciones respecto a la descomposición permanente-transitorio (P-T) de las series son:

- Se introduce una nueva definición de descomposición basada en la definición de Quah (1992), y que introduce la restricción de que el componente permanente este sólo formado por perturbaciones permanentes.
- Basados en los modelos ARTIMA, se proponen dos métodos no lineales de descomposición permanente-transitorio. El primero se obtiene de la aplicación de la nueva definición de descomposición P-T. El segundo se obtiene a partir de la definición propuesta por Beveridge y Nelson (1981), (B-N en adelante).
- Comparada con la descomposición B-N lineal, la descomposición B-N no lineal basada en los modelos ARTIMA, elimina la restricción de que ambos componentes estén perfectamente correlados. Con relación a la descomposición $UC - 0$ propuesta por Harvey (1985), Clark (1987) y Pierce (1979) entre otros, comparte la propiedad de que los componentes pueden ser incorrelados o parcialmente correlados, pero elimina la restricción de información a priori. En el Capítulo 4 se prueba que el comportamiento del componente transitorio en la descomposición no lineal B-N se encuentra siempre entre ambos, el del B-N lineal y el del $UC - 0$.
- En general, en la medida que el contraste de hipótesis en los modelos ARTIMA sea factible, los criterios en los que se basan las nuevas descomposiciones serán también

contrastables, permitiendo una endogeneidad en el método de descomposición no alcanzado por ninguna metodología propuesta hasta ahora.

Las líneas de investigación que este trabajo abre son muy numerosas, en principio, centradas en la extensión multivariante de los métodos propuestos. Entre ellas, caben destacar dos:

- La extensión a los modelos VARMA. En concreto, se tratará de analizar bajo que condiciones las asimetrías en la persistencia eliminan el problema de la identificación de las perturbaciones en los modelos VARMA (ver Quah, 1992).
- Relacionado con el marco anterior, se encuentra la posibilidad de definición de nuevas relaciones de equilibrio a largo plazo o cointegración. En el contexto de un modelo de corrección del error, una ecuación de cointegración modelizada a través de un modelo ARTIMA, permitirá extender la definición de cointegración parcial al caso en el que solo bajo determinados regímenes las series están cointegradas. La definición actual de cointegración parcial hace referencia a los casos en los que existen periodos de no atracción entre las variables, pero en el que las series están cointegradas. Esta definición más general de cointegración parcial permite que la relación de equilibrio a largo plazo entre las variables no sea constante, si no que pueda cambiar bajo determinadas condiciones económicas.

Validez asintótica de los métodos propuestos

Como hemos visto, una de las principales aportaciones de los métodos de identificación propuestos reside en la posibilidad de contrastar la validez de los criterios utilizados. Para ello son necesarios varios resultados respecto a la estimación de los modelos. En primer lugar, los parámetros del modelo han de estar identificados bajo el método de estimación propuesto, que a su vez, ha de ser un método consistente. En segundo lugar, se han de

proponer estrategias de contraste de las hipótesis básicas de modelo, relacionadas con los criterios de identificación, para las que se han de obtener resultados de validez asintótica. Estas hipótesis son fundamentalmente dos: la existencia de asimetrías en la persistencia, y cuando estas existan, su extensión al caso de transitorio y permanente. Todos estos resultados se presentan en los Capítulos 2, 3 y 4. Para ello, se han distinguido dos casos fundamentales según la variable threshold sea observable, ARTIMA-Observable en el Capítulo 2, o la propia perturbación, ARTIMA-Shock en el Capítulo 3. A su vez, en los modelos ARTIMA-Shock se analizan los casos particulares del signo y el tamaño, ARTIMA-Shock-Sign y ARTIMA-Shock-Size respectivamente. A continuación presentamos un resumen de los resultados:

- En un paso inicial, se obtienen las condiciones suficientes para la existencia de perturbaciones transitorias y permanentes bajo la modelización ARTIMA, así como las de invertibilidad, ya que se trata de modelos no lineales y con medias móviles. La invertibilidad es una condición necesaria para la estimación consistente de las perturbaciones, que permite la predicción y que dota de sentido a la definición de los conjuntos de información basados en las propias perturbaciones.
- El método de estimación propuesto para ambos modelos se basa en mínimos cuadrados. Para este método se obtienen las condiciones suficientes para la identificabilidad de los parámetros y consistencia de los estimadores. Los resultados obtenidos son diferentes según la variable threshold sea observable o la propia perturbación. En el primer caso, los resultados replican los obtenidos en el marco de los modelos TAR, tasa de convergencia $T^{1/2}$ en los estimadores “continuos” y superconsistencia del estimador del parámetro threshold. Para el caso ARTIMA-Shock, debemos distinguir, a su vez, dos casos. En el caso del signo, con $r^0 = 0$ conocido, los resultados son los mismos que para el ARTIMA-Observable, dado que el modelo será un threshold continuo. En el otro caso, $r^0 \neq 0$, al igual que en el caso del tamaño, todos los estima-

dores serán superconsistentes. La condición $r^0 \neq 0$, garantiza la discontinuidad del modelo. Este es un resultado novedoso y que se deriva de que todos los parámetros comparten la discontinuidad típica del parámetro threshold en el modelo TAR.

- Respecto al contraste de existencia de threshold o asimetrías, cuando r^0 es conocido, en el caso observable, e igual a 0 en el caso ARTIMA-Sign, se propone un contraste de Wald para la igualdad de los parámetros en la media móvil. Su distribución asintótica bajo la nula de simetría es una χ^2_1 . Para el resto de casos, se proponen dos tipos de estadísticos de tipo supremo, dado que en ambos, bajo la hipótesis nula de simetría, el parámetro threshold no está identificado. El primero, se basa en la estimación del modelo bajo la hipótesis alternativa, Capítulos 2 y 3, mientras que el segundo, en su estimación bajo la nula, Capítulo 4. Este último es más fácil de calcular, aunque con peor potencia. En ambos casos se obtiene la distribución asintótica del estadístico, que dependerá de parámetros ruidosos y no se podrá tabular. Por ello, se proponen aproximaciones bootstrap para obtener los p – valores.
- Cuando se rechace la hipótesis nula de no existencia de asimetrías, se propone el contraste de existencia de perturbaciones transitorias y permanentes. Éste, se traduce en un contraste de raíz unitaria en algún régimen de la media móvil. Bajo determinadas condiciones, se obtendrá la normalidad asintótica del estadístico propuesto. En el caso del ARTIMA-Shock discontinuo, este resultado se obtiene después de una estimación en dos etapas de modelo. La estimación de la segunda etapa permite evitar la discontinuidad de los estimadores presente en la primera etapa, aprovechando la superconsistencia de los mismos.
- En el caso del modelo ARTIMA-Shock para el signo, el contraste de existencia de threshold propuesto en la tesis no es nuevo, ver Brännas y Gooijer (1994) o Elwood (1998). Sin embargo, en ninguno de los casos anteriores se obtuvieron resultados de

validez asintótica, sí logrados, por el contrario, en esta tesis.

Las líneas de investigación futuras abiertas en este campo son también muy numerosas. Fundamentalmente, nos centraremos en las siguientes:

- Validación teórica de los métodos bootstrap propuestos en la Tesis para la obtención de los p-valores en los contrastes de existencia de efecto threshold, la cual se obtendría siguiendo los pasos propuestos por Hansen (1996).
- Extensión de los resultados de validez asintótica a los casos de innovaciones que no sean *iid*, fundamentalmente aquellos en los que éstas presentan heterocedasticidad condicional. Otra extensión muy relevante, relacionada con las condiciones suficientes, es la que permita reducir las condiciones de dependencia de la variable threshold, posiblemente hasta $L_2 - NED$ (ver Pötscher y Prucha, 1997).
- Consideración de modelos ARTIMA más flexibles. En concreto, se consideraría la posibilidad de efectos threshold, no solo en la media móvil, si no también en la constante. Esto permitiría aumentar, bajo determinadas condiciones, el número de casos en los que es posible generar perturbaciones transitorias y permanentes. En particular, el caso del signo.
- La extensión de los resultados a modelos no lineales continuos, que podría ayudar a relajar las condiciones suficientes para la invertibilidad de los modelos y la validez asintótica de los métodos propuestos.
- Desarrollo de un contraste para la hipótesis nula de $r^0 = 0$ en el caso de los modelos ARTIMA-Shock.
- Finalmente, desarrollo de un contraste conjunto para la hipótesis, $H_0 : \theta_1 \neq \theta_2 = 1$.

Larga memoria

En muchas series temporales, y en particular en las económicas, no siempre es factible aceptar la hipótesis nula de que el proceso es $I(0)$ o $I(1)$. Los procesos ARFIMA permiten una solución intermedia que implica que todas las perturbaciones tienen la misma persistencia, que no es ni la permanente de un $I(1)$, ni la transitoria de un $I(0)$, sino algo intermedio. Los modelos ARTIMA proponen una solución alternativa que implica que algunas perturbaciones son transitorias y otras permanentes. Así, en la Tesis se realiza con un estudio de la relación existente entre los modelos threshold y las propiedades de larga memoria de los procesos ARFIMA. Concretamente, analizaremos los modelos ARTIMA-Shock-Size, introducidos en la Tesis, y los Threshold Unit Root, TUR. Esta relación es muy importante dada la escasez de modelos económicos que justifiquen la aparición de este tipo de propiedades teóricas. Siguiendo los resultados obtenidos previamente por otros autores en referencia a otro tipo de modelos no lineales, como son los modelos de breaks aleatorios en la media o el modelo STOPBREAK de Engle y Smith (1999), en el Capítulo 5 se obtienen los siguientes resultados que se resumen a continuación:

- Se prueba que cuando el parámetro threshold depende del tamaño muestral, los modelos threshold, TUR o ARTIMA, son capaces de reproducir el comportamiento asintótico de la suma parcial de un proceso ARFIMA. También se justifica como la agregación sectorial de procesos TUR puede generar series con las propiedades de larga memoria de un proceso ARFIMA, replicando los resultados de Granger (1980) y Lippi y Zaffaroni (1999) para los $AR(1)$. Sin embargo, este resultado no es extensible a los modelos ARTIMA, cuya agregación dará lugar a procesos $I(1)$.
- A través de un estudio de Monte Carlo, se lleva a cabo un análisis detallado del comportamiento en muestras finitas de los contrastes de larga memoria propuestos por Geweke y Porter-Hudak (GPH) (1983) y Tanaka (1999) en los modelos ARTIMA

y TUR. Nos centraremos, fundamentalmente, en el porcentaje de rechazos de las hipótesis nulas $I(0)$ e $I(1)$. Cuando el porcentaje de rechazos de ambas hipótesis sea elevado, también lo será la probabilidad de confundir la serie con un proceso ARFIMA. En general, los resultados son consistentes con la teoría desarrollada a lo largo del capítulo ya que para diferentes tamaños muestrales y valores del parámetro threshold la confusión es muy probable. Para comprobar si esta confusión es exclusiva del carácter no lineal del modelo, los resultados se comparan con los obtenidos para la mejor aproximación ARIMA y ARFIMA de los modelos ARTIMA y TUR. En general, el comportamiento es muy similar al obtenido con su aproximación lineal ARIMA y ARFIMA. El comportamiento más desigual se obtiene para el caso de los modelos TUR y en el contraste de Tanaka. Como cabía esperar, el contraste de Tanaka resulta más robusto que el de GPH, dado su carácter paramétrico. Pero por lo mismo, su éxito se basa en la correcta especificación del modelo.

- Finalmente, también se analiza el comportamiento del contraste en muestras finitas de Dickey-Fuller aumentado, a través de un estudio de Monte Carlo. En general, para evitar la distorsión en el tamaño, caso ARTIMA, se requerirá un elevado número de retardos en la especificación del contraste. El efecto en la potencia, caso TUR, resulta inferior.

Las posibles líneas de investigación futuras, se centrarán fundamentalmente en dos campos:

- Extensión del estudio de Monte Carlo a nuevos contrastes desarrollados recientemente y que presentan una mayor robustez a breaks en la media.
- Nuevos modelos no lineales que generen larga memoria, y que permitan asimetrías en la persistencia.

6.1. Bibliografía

Beveridge, S. y Nelson, C.R., (1981): “A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘Business Cycle’”. *Journal of Monetary Economics*, Vol. **7**, 151-174.

Brännas, K. y Gooijer, J.G.D. (1994): “Autoregressive-asymmetric moving average models for business cycle data”. *Journal of Forecasting*, Vol. **13**, 529-544.

Burns, A.F. y Mitchell, W.C. (1946): *Measuring Business Cycles*, Columbia University Press.

Clark, P.K. y (1987): “The Cyclical Component of U.S. Economic Activity”. *Journal of Quarterly Journal of Economics*, Vol. **102**, 797-814.

Elwood, S.K. (1998): “Is the persistence of shocks to output asymmetric?”. *Journal of Monetary Economics*, Vol. **41**, 411-426.

Engle, R.F. y Smith, A.D. (1999): “Stochastic Permanent Breaks”. *The Review of Economics and Statistics*, Vol. **81**, 553-574.

Geweke J. y Porter-Hudak S. (1983): “The estimation and application of long memory time series models”. *Journal of Time Series Analysis*, Vol **4**, N4, 221-238.

Hansen, B.E. (1996): “Inference when a nuisance parameter is not identified under the null hypothesis”. *Econometrica*, Vol. **64**, 413-430.

Harvey, A.C., (1985): “Trends and Cycles in Macroeconomics Time Series”. *Journal of Business and Economic Statistics*, Vol. **3**, 216-27.

Pierce, D.A., (1979): “Signal Extraction Error in Nonstationary Time Series”. *The Annals of Statistics*, Vol. **7**, 1303-1320.

Pötscher, B.M. y Prucha, I.R. (1997): *Dynamic Nonlinear Econometric Models: Asymptotic Theory*, Springer-Verlag Berlin and Heidelberg GmbH & Co. K.

Quah, D., (1992): “The relative importance of permanent and transitory components: identification and some theoretical bounds”. *Econometrica*, Vol. **60**, 107-118.

Tanaka, K. (1999): “The nonstationary fractional unit root”. *Econometric Theory*, Vol. **15**, N 4.